Discourse, justification and education: Jurgen Habermas on moral epistemology and dialogical conditions of moral justification and rightness

Abstract. In this paper Walter Okshevsky addresses the question of whether a certain form of dialogically derived agreement can function as an epistemic (universal and necessary) criterion of moral judgement and ground of moral authority. Okshevsky examines arguments for and against in the literature of educational philosophy and develops Jurgen Habermas’s affirmative answer as presented in his Discourse Theory of Morality. Habermas’s position is articulated as a moral epistemology (“strong dialogicality”) and is developed through his critique of the “monologism” of certain aspects of Kant’s moral theory. Okshevsky concludes with a consideration of some educational implications of Habermas’s position.

He who knows only his own side of the case, knows little of that. His reasons may be good, and no one may have been able to refute them. But if he is equally unable to refute the reasons on the opposite side; if he does not so much as know what they are, he has no ground from preferring either opinion. … Nor is it enough that he should hear the arguments of adversaries from his own teachers, presented as they state them, and accompanied by what they offer as refutations. That is not the way to do justice to arguments, or to bring them in real contact with his own mind. He must be able to hear them from persons who actually believe them; who defend them in earnest, and do their utmost for them.

John Stuart Mill, *On liberty*

Introduction

An important epistemological problem underlying many questions regarding the nature, conditions and learning of moral deliberation and judgement is whether agreement on the rational justifiability of a norm or principle constitutes an *epistemic* criterion of moral truth or rightness. Can agreement, attained under certain conditions, play a necessary and/or sufficient role in providing criteria for the objective justification of moral rightness claims? In other words, could such agreement possess epistemic status in moral deliberation or do we require criteria independent of and external to agreement in order to assess the justifiability of the truth or rightness of any attained agreement itself? Should agreement in some form turn out to possess epistemic status, the ideal of such agreement would be understood to permit agreed upon norms, principles and judgements to transcend forces of ideology, power and bias more effectively than is
possible for any individual thinker relying solely on her own rational and affective capacities for deliberation and judgement. On the strong version of the affirmative thesis held by Jurgen Habermas and others – a version I will term “strong dialogicality” - agreement satisfying certain conditions of dialogical symmetry and reciprocity is understood to be epistemic in these terms as well as to constitute a particular kind of reason of its own. More specifically, agreement attained in “discourse” (communicatively-oriented argumentation) under satisfaction of necessary procedural conditions is understood to be constitutive of moral justification and, indeed, moral rightness itself. Habermas’s claim is that such a dialogically-based conception of justification and rightness provides the most appropriate grounds of moral authority given the obligations of a pluralist democracy in a post-metaphysical age. Should Habermas’s epistemological argument be sound, it would comprise an important resource and ground for an impartial articulation and justification of educational aims and practices in the moral domain that avoid substantive culturally-relative conceptions of what is good or worthwhile to learn. Should a form of agreement turn out to be an epistemic criterion of normative justification, the educational aim of fostering requisite dialogical abilities and corresponding motivations to search for epistemically valid agreement with others would yield essential educational imperatives of developing cognitive, affective and conative abilities and dispositions necessary for competent engagement in discourse as a particular form of argumentation. In this way, epistemological considerations would provide cogent justification for specific educational interventions and outcomes. Habermas writes:

[T]he learning processes by which we acquire theoretical knowledge and moral insight, revise and extend our evaluative language, overcome self-deceptions and difficulties of understanding, depend on [discursive] argumentative practices.1

[L]earning in the moral domain is assessed in terms of how inclusive … a consensus reached through reason giving is (emph. deleted).2

In this paper, I reconstruct the moral epistemology developed by Habermas’s Discourse Theory of Morality focussing on its claim that a specific form of ideal agreement, i.e., mutual and joint agreement attained in argumentation under satisfaction
of procedural conditions of symmetry and reciprocity, constitutes a necessary and universally valid epistemic criterion for moral justification: “Moral justifications are dependent on argumentation actually being carried out.”3 The validity or rational acceptability of statements “… can be established only through discursive engagement using available reasons” (TJ, 247). As the simple fact of agreement attained by non-discursive means clearly bears no necessary epistemic warrant, I shall have little to say about it. Important to emphasize at the outset is that Habermas acknowledges that we agree on P because we take P to be true or right on the basis of reasons; P is not true or right in virtue of simple de facto agreement on it. Agreement on a norm, like any belief or judgement, is fallible. The ideal agreement that functions as an epistemic requirement for Habermas is a discursively derived joint and mutual agreement attained under precise procedural conditions expressive of a universal egalitarianism. Good reasons for moral beliefs and actions are to be had only through discourse as a distinct procedure of argumentation. Habermas maintains that discursive agreement functions not only as an epistemic criterion but also serves as the ground of the only legitimate conception of moral rightness and authority possible in a postmetaphysical age where impartial and universally valid acceptance of principles grounded in substantive cultural and religious values is no longer deemed a viable option. The Discourse Theory’s project of articulating and justifying epistemic grounds of moral norms maintains that only norms that are or could be agreed to within procedurally valid discursive examination may justifiably claim moral rightness and be deemed morally right.

My examination will involve a consideration of arguments presented by a number of educational and general philosophers who have addressed this problematic. Relative to these past analyses, I believe we can offer here a more comprehensive and systematic consideration of the philosophical origins and grounds of Habermas’s position and his Discourse Theory. I want to as well illustrate Habermas’s deployment of “strong dialogicality” in his critique of the “monological” character of Kant’s conception of the Categorical Imperative and I present his re-interpretation of this Imperative along dialogical/discursive lines. I conclude with a brief consideration of the educational implications of Habermas’s moral epistemology.
I

Inspired by Habermas’s Discourse Theory, Ken Strike and Jonas Soltis pose the question of whether the rightness of ethical judgements and principles is epistemically tied to agreement on their acceptability. They recognize with Habermas that, should this turn out to be the case, not just any kind of de facto agreement reached through dialogue could be said to contribute to the justification and rightness of ethical principles and judgements. In order for agreement and procedures of dialogue to either contribute to, or be somehow constitutive of, justified claims to moral rightness, the agreement must be the outcome of a dialogue satisfying certain procedural conditions (ET, 125).

Following Habermas, they offer the following ideal features as necessary conditions for establishment of relations of symmetry and reciprocity between interlocutors as required by discourse: 1) dialogue must be free of coercion and deception, 2) perspectives of all relevant parties must be included, 3) deliberations and decisions arrived at must be public, with no relevant considerations suppressed and no argument excluded, 4) all parties are to be treated as equals and their respective interests equally and fairly respected. As well, dialogue must meet the condition of reciprocity:

Individual participants in the decision should be able and willing to project themselves into the perspective of other parties in the discussion and to find any decision reasonable from the variety of available perspectives, not only from their own” (ET, 125).

(See “Conclusion” below for an amplified statement of these conditions.) Strike and Soltis hold that these conditions are necessary for an “open and undominated dialogue” and insofar as procedural conditions are adequately satisfied it is possible that dialogue may at times serve an epistemic function in claims to moral rightness: “That an ethical decision results from an open and undominated discussion may be a factor that actually contributes to its being a right decision” (ET, 125). What remains unclear in their position is whether the “contribution” that agreement on the acceptability of a decision under satisfactory fulfilment of these dialogical conditions makes to rightness claims constitutes a necessary and universal, i.e., epistemic, condition of ethical justification. Is the agreement that they say may be a contributing factor to correctness to be understood in
the sense of Habermas’s thesis of “strong dialogicality” – i.e., that independent of satisfactory agreement under discursive conditions, no judgment can justifiably claim moral rightness and be right? Their concluding claim that agreement “may be a factor that actually contributes to its being a right decision” leaves this an open question. Their view would as such seem to allow for the possibility that there are other norms or factors independent of conditions of dialogue that can play an equivalent epistemic role such that a judgement arrived at through a closed and dominated dialogue may yet be right and/or be justified in virtue of satisfying such other independent criteria. That view would imply that any attained agreement must itself be epistemically assessed by means of these independent criteria. This would then be a position we can call “weak dialogicality.” This view allows that an individual thinker or a deliberating group could come up with a justifiable or right decision without satisfying conditions of an open and undominated dialogue, indeed without engaging in discourse at all. Individual moral agents would as such be deemed capable of making justified moral claims on their own, independent of the kinds of collective reflection, argument and inquiry required by dialogical conditions. A principle such as Kant’s Principle of Equal Respect, accepted by Strike and Soltis as a defensible moral principle and believed by them to be embedded in the identified dialogical conditions, would in that case constitute an epistemic criterion serving as an external standard of assessment of any discursively-attained agreement. As Kant himself believed, the principle serves as an objective, “Categorical,” standard for the assessment of the justifiability of any given outcome of deliberation on the moral permissibility of maxims or principles, and one which each of us on our own can cogently apply. That Strike and Soltis view Habermas’s conditions of discourse to meet one of Kant’s formulations of the Categorical Imperative thereby suggests again that they indeed hold a position of weak dialogicality.

Mark Weinstein takes such procedural conditions as equality of access to and equality of participation in dialogue to constitute “the social core of Habermas’s ideal speech situation.” Weinstein maintains that not only moral failure but epistemic failure is displayed in the failure to ensure inclusion of different perspectives, especially those held by individuals and groups traditionally marginalized, silenced or oppressed:
“Persistent disregard of blatant injustice, injustice that should be readily perceived from within the practice that disregards them, thus becomes an essential clue to the epistemological failings of the practice.” Procedural violations of dialogical norms such as exclusion of perspectives relevant to deliberation on the moral justifiability of a norm fail to recognize that such inclusion is “dialectically invaluable” and as such undermine the justifiability of individual moral judgement or any collective agreement achieved by groups of deliberators. Such transgression for Weinstein is targeted not only on moral grounds but also on epistemic grounds: “the systematic exclusion of a point of view indicates a structural failing in the discourse frame” (RH, 33-34). Procedural conditions identified by Habermas’s “ideal speech situation” are epistemically necessary in that they enable and condition the rational justification of norms. Sharon Bailin calls Weinstein’s view here “the strong version” of the role of procedural conditions such as inclusion in differentiation from its weaker version – a distinction we have marked as one between “strong dialogicality” and “weak dialogicality.” While Weinstein does not always clearly present his position with careful attention to this distinction of which he approves, I believe Sharon Bailin accurately renders his considered position:

[For Weinstein] … some discourse frames are epistemically inadequate by virtue of their exclusionary nature. … The argument here is not that inclusive frames necessarily produce theories [and discourses] which are more adequate according to independent epistemic criteria, but rather that inclusivity is, itself, a criterion of epistemic worth. … It is not simply [as per the weak version] that the disregard of certain perspectives makes a discourse frame less likely to produce theories which are more worthy by other standards. Rather the fact that a frame disregards certain perspectives despite their apparent availability constitutes an epistemological failing in that this reveals a pathology in the frame [itself].

R. Jay Kilby, in his critical reply to Harvey Siegel on the question of the epistemic status of inclusion, also appeals to Habermas’s thesis of strong dialogicality: “A communicative theory of rationality holds the promise of … establishing that epistemic virtues such as inclusion are inherent to rationality rather than dependent upon [external and independent] universal criteria [for their justification]”8. In differentiation from Weinstein’s position, however, but in keeping with Habermas’s own view, Kilby
does not believe procedural conditions themselves comprise substantive moral norms. As we shall see below, this comprises a central point of contention between Kilby and Siegel. Procedural conditions defining the ideal speech situation are understood to function as epistemic criteria for the assessment of the justifiability of moral norms and principles.

[I]nclusive discourse and absence of force constitute an ideal speech situation against which actual attempts to offer moral justification for acts must be measured. In this respect, inclusion and related epistemic virtues are necessary components of practical reason (CT, 304).

Like Weinstein, Kilby’s language at times slides into ambiguity between strong dialogicality and the weak version that we will see Siegel and Rescher to happily allow. Nevertheless, he develops a number of cases in which the procedural condition of inclusion of different perspectives is illustrated to function as an epistemic criterion and he provides important discussions of the nature of epistemic criteria in criticism of a number of Siegel’s views. Kilby’s argument that a reliance on epistemic virtues and dialogical procedure are necessary ingredients of moral justification is particularly compelling in the case of disagreement on which criteria should be used in justifying belief and what kind of evidence is most relevant in justifying particular kinds of belief.9

There is no ambiguity or ambivalence in Harvey Siegel’s position on our question. He explicitly rejects Habermas’s thesis of strong dialogicality. Siegel is willing to acknowledge, as per weak dialogicality, that engagement in Habermasian dialogue can at times help to get things right. On this he is in agreement with Strike and Soltis. But he wants to emphasize that agreements issuing from dialogue, however procedurally correct, may yet get things wrong. Agreement on the acceptability of a moral claim established under dialogical conditions in itself bears no necessary epistemic status or warrant for Siegel. In granting the weak version of dialogicality he only acknowledges that procedural failure in dialogue could result in epistemological failure and so we are well served by cultivating the epistemic virtues.10
Inclusion may well be an epistemological (methodological) virtue in the sense that the more voices and perspectives are included in theorizing [or deliberation], the more likely it is that epistemically worthy beliefs/hypotheses/theories will be generated and accepted (italics omitted).

Nevertheless, provision by the virtues of such a service does not entail any necessary epistemic worth. The connection between, on the one hand, an epistemic virtue such as open-mindedness or ability and willingness to identify one’s prejudices and preconceptions within deliberation and, on the other, an epistemic criterion by which we assess warrant proffered for a claim is a contingent one: it bears no necessary conclusions regarding epistemic status or warrant of dialogical agreement. Any de facto agreement, like any belief held by an individual, remains open to assessment on the basis of objective principles and criteria of justification that must logically be external to and independent of belief or attained agreement, however comprehensively inclusive and procedurally virtuous agreement on the truth or rightness of belief may be. Procedural conditions of dialogue are understood by Siegel to comprise moral conditions, more specifically, conditions of “procedural justice.”

This immediately brings to mind Habermas’s views of ideal communicative situations, which require “equitable access to the dialogue and equality within it.” When [discourse] frames silence or marginalize, they fail to meet Habermas’ criteria; in so failing, they fail to treat potential participants in dialogue justly (RRF, 133).

But the just treatment of interlocutors in dialogue is deemed to comprise a necessary epistemic requirement neither for the justification of claims to moral rightness nor for the meaning of moral rightness itself. With specific reference to the procedural condition of inclusion, Siegel writes:

A dialogue in which all have full access and in which all are treated equally, with respect, may nevertheless result in agreement on beliefs which are false or unjustified. On the other hand, a dialogue to which certain people are excluded may nevertheless result in beliefs which are true or justified (RRF, 133).

Siegel would thus concur with Strike and Soltis that the identified conditions of the “ideal speech situation” or an “open and undominated dialogue” seek to secure the
moral treatment of dialogical interlocutors. However, the ambiguity we noted in their position is not present in Siegel’s as he clearly discounts the possibility of any sound inference from satisfactory procedural fulfilment of dialogical conditions to the justifiability of belief or judgement. Procedural conditions do not constitute necessary conditions for the “epistemic worthiness” of beliefs, judgements or agreements:

[E]pistemic considerations … [refer to] the way in which beliefs, judgments and actions are reasonable (or not) because of their substantive, contentful relations to putative reasons which support them (or not) (RRF, 107).

Consequently, procedural failure to abide by dialogical conditions counts as a moral failure and not necessarily an epistemic failure. As with Strike and Soltis, Siegel understands such failure as a violation of Kant’s Principle of Equal Respect for Persons, and he seems to agree with Kant on the status of this principle within sound moral reasoning. For Kant, as we shall see below, this is a principle the objectivity and universality of which does not itself require dialogical agreement to possess such status. It is therefore able to be deployed objectively as an external and independent epistemic criterion in assessing the rational acceptability of a moral judgement or the universalizability of a maxim. Given an understanding of the principle in such terms, it follows that moral failures in dialogical procedure are not necessarily epistemic failures precisely because we may be justified in our beliefs despite failing to adequately satisfy one or more dialogically required procedural conditions such as inclusion, and despite having arrived at our beliefs independent of engagement in discursively governed dialogue. (We will see more of Kant’s “monologism” below.) Siegel warns against erroneously conflating rules governing the procedure or conduct of dialogue or inquiry with criteria for the assessment of the products of dialogue and inquiry (ibid). To do so is to fail to respect the distinction between the rationality of a procedure or process and the rationality of the products of a procedure or process.12 The epistemic status of belief – i.e., its truth or rightness and the “evidential or probative support” brought forth for its justifiability – is logically separate from questions regarding the processes leading to the formation of belief (MR, 214). For Siegel, how we reach agreement on some issue, together with whether we reach agreement on the issue, are matters bearing no necessary
relation to the soundness of arguments and the epistemic worthiness of the warrant provided in justification of beliefs and judgements. Siegel concludes:

[I]nclusion is not a criterion by which to evaluate epistemic worthiness. We cannot say of a candidate belief/hypothesis/theory that it is true (false), or more (less) probable, or enjoys a higher- (lower-) than-otherwise justificatory status, because it was generated in inclusive (exclusive) circumstances (IEV, 97).

Nicholas Rescher takes a very similar view in his support of weak dialogicality and rejection of the strong thesis of dialogicality. While he grants that consensus can at times function to check for error, bias or carelessness, he maintains that consensus is neither a necessary nor a sufficient means for attaining rational ends of either agency or inquiry.

In the circumstances in which we labour in this world, consensuality is neither a requisite for, nor a consequence of, rationality in the conduct of inquiry – it is neither a necessary nor sufficient condition for it.13

Like Siegel, Rescher maintains that the regulative principles of dialogical procedure identified by Habermas clearly allow for the possibility that satisfactory accordance with them may yet yield mistaken results (P, 13). The kind of consensus that rationality requires, when it requires it at all, is grounded in sound warrant or evidence for belief. Good reasons rather than an attained consensus is what genuinely counts in justification. Any consensus achieved remains open to critical assessment of the rationality of the grounds proffered for that agreement by epistemic criteria independent of and external to agreement. In Rescher’s view, Habermas over-values the epistemic worth of agreement and fails to recognize that a product of inquiry or deliberation may be true, right or rational regardless of whether a consensus has been achieved on it (P, 13-16).

In the following sections, I address Strike and Soltis’s ambivalence and Siegel’s and Rescher’s objections to strong dialogicality by examining the grounds of Habermas’s position and I will illustrate the position in its application to criticisms Habermas makes of a number of central features of Kant’s moral theory. While both Weinstein and Kilby
subscribe to Habermas’s thesis of strong dialogicality, and offer important arguments that rely on the thesis as a premise, their explications of Habermas’s arguments for this position tend to be skeletal and do not situate his justification of it within the broader Discourse Theory. As a result, their considerations of the grounds of the thesis are not developed sufficiently for a systematic and comprehensive reply to objections of the kind raised by Siegel and Rescher among others. The following sections attempt such a more comprehensive reply.

One important aspect of Habermas’s moral epistemology requires to be clarified at the beginning. While discursive agreement on the rational justification of a moral norm or judgement functions as an epistemic criterion, it does so in a manner different from other criteria such as, for example, the logical validity of an argument. The validity of arguments is such that the criterion of validity is either fully met or not met at all in each case of inference. Procedural conditions of valid discursive agreement, however, constitute conditions that are “simultaneously counterfactual and practically efficacious.”14 While indispensable within rational moral deliberation and judgement, they cannot ever be fully satisfied by any actual agreement on the rightness of a norm. Epistemically valid moral agreement is hence qualitative, not binary; it comprises a scalar criterion satisfaction of which can admit only of degrees. We will see Habermas’s position to counter Siegel’s view that the soundness of a moral argument depends ultimately on validity and truth of premises with the claim that judgement and agreement on the rational acceptability of a moral norm must be considered as “more or less justified,” and this depending on the degree to which procedural conditions have been met. (This is of course not to say that Siegel’s sole criterion for justification in general is deductive validity and soundness.) The extent to which these procedural conditions have been met determines the degree to which the norm or judgement is epistemically warranted. In Siegel’s lexicon, we can say that for Habermas it is precisely the degree of fulfilment of discursive procedural conditions that establishes “higher- (lower-) than – otherwise justificatory status.” It is in such a form that the epistemic force of discursive agreement functions as a necessary regulative ideal both structuring the procedure of discourse and governing epistemic assessment of the outcomes of discourse. I will
attempt to establish that the scalar nature of dialogical agreement in moral justification is not a mark of its inadequacy or incoherence as a genuine epistemic criterion in the moral domain but comprises rather a necessary feature of the only kind of assessment of the impartiality of moral judgement and deliberation available to us in the postmodern world. Strong dialogicality is as such a form of Constructivism – an anti-realist epistemology denying the existence of facts, reasons or entities existing independent of and external to reasoning agents, specifically reasoning agents under discursive conditions of symmetry and reciprocity. Pace the kinds of objections proffered by Siegel and Rescher, Habermas maintains there is no cogent view of moral justification or the rightness of moral judgement that is ultimately independent of deliberating agents collectively engaged in argumentation governed by ideal conditions of discourse.

[Discursive] procedures and reasons are so closely interwoven with each other that there cannot be any evidence or criteria of assessment that are completely prior to argumentation [and] that do not have to be justified in turn in argumentation and validated by rationally motivated agreement reached in discourse under the presuppositions of argumentation. … Because there cannot be “ultimate” evidence or “decisive” arguments in dealing with substantive questions, we must appeal to the pragmatics of the procedure of argumentation in order to explain why we can even think ourselves capable of raising and redeeming context-transcending validity claims.15

Once properly understood, writers such as Siegel and Rescher have nothing to fear from strong dialogicality as a moral epistemology. The soundness of moral arguments and the goodness of reasons are not abandoned by Habermas. However, the origins and means of assessment of such arguments, together with the formulation of necessary ideal conditions of justification, require to be recognized as dialogica in character.

II

Habermas’s conception of discourse is a distinct and univocal one. As a term of art, “discourse” refers to a form of dialogue and argumentation involving persons and parties as interlocutors engaged in the construction of a norm or policy that is generalizable or “universalizable” in Kant’s lexicon. Such communication intends the construction of a norm equally in the interests of all actually or potentially affected by it’s implementation. Strong dialogicality states that no norm may justifiably be accepted as
universalizable, and hence morally justified and right, independent of engagement in dialogue with others. For it is only through such collective engagement that participants are given the opportunity to present to others perspectives, needs and interests relevant for the construction of a universalizable norm. What is sought in discourse is an impartial judgement or decision from the epistemic perspective of universal egalitarianism – a perspective defined by formal procedural conditions ideally ensuring symmetry and reciprocity across interlocutors’ contributions to the process of argumentation. Discourse is possible only on condition that interlocutors remain convinced that all are abiding by procedural conditions of symmetry and reciprocity. Should one or more conditions be claimed to be abrogated, discourse is either terminated or continues on, now with an examination of the validity of the claim to violation. Any judgement or decision discursively arrived at must have attained mutual and joint agreement and be the result solely of argumentation, understood as the collective giving and assessing of reasons:

Participants in argumentation have to presuppose in general that the structure of their communication … excludes all force – whether it arises from within the process of reaching understanding [and agreement] itself or influences it from the outside – except the force of the better argument (and thus that it excludes, on their part, all motives except that of a cooperative search for the truth) (TCA 1, 25)

Opponents of strong dialogicality may rest assured that good reasons and justifiable moral judgements are here internally connected: “[t]o say that I ought to do something means that I have good reasons for doing it” (MCCA, 49). As is the case with all (moral) judgement and argument, any discursively attained agreement on a norm is fallible, regardless of the scope of agreement. A norm or policy accepted as one representing a generalizable interest may subsequently come to be judged invalid, and thus morally wrong, upon disclosure of intentional or unintentional violation of one or more procedural conditions (TJ, 108). Fallibility clearly extends to interlocutors’ abilities to detect procedural violations. But fallibility in applying an epistemic criterion is not itself a necessary or conclusive mark against its suitability as such a criterion. The appropriateness or necessity of the criterion of soundness of argument, for example, is not impugned by cases of invalid reasoning or the making of false claims. It is simply violated, typically unintentionally. Habermas writes: “[T]he rational acceptability of a
statement ultimately rests on reasons in conjunction with specific features of the process of argumentation itself” (GA, 44).

Let us now take a closer look at the thesis of strong dialogicality itself. Why does Habermas maintain that valid (mutual and joint) agreement under procedural conditions of symmetry and reciprocity, understood as “idealized counterfactuals,” constitutes an epistemic criterion, and the only cogent criterion, of moral justification such that a norm or judgement developed in violation of, or even independence from, satisfactory fulfilment of these dialogical conditions fails to be properly justified? I believe the answer rests primarily in Habermas’s view of the character of justifiable moral authority along with his conception of the distinct and irreplaceable functions which moral argument serves in a pluralist social order. Fundamental here is his view that the enterprise of morality is not accurately characterized as a primarily theoretical form of knowledge and inquiry. This is to say that the practice of morality cannot be pursued within a strictly propositional attitude for which meaning and truth of judgement are decided “objectively” via truth conditions for statements. Habermas’s dialogical approach counters the comprehensiveness and adequacy of such a semantic conception of meaning and truth. Building on Speech Act theory as developed by John Austin and John Searle, Habermas understands a moral claim as a distinct kind of speech act used by a speaker to publicly state or assert a judgement (“validity claim”) and, if necessary, be willing to secure others’ acceptance of it in collective interaction through the discursive giving and assessing of reasons. Central to the functions of moral judgement is the coordination and regulation of social interaction and, as circumstances require, the restoration of harmonious integration of ends through conflict resolution. No social order can exist without social cooperation between individual agents and the continued reinforcement of agents’ mutual expectations of social coordination of means-ends activities. Successful attainment of ends and interests requires the intersubjective recognition of the stability of an order able to effectively integrate individuals’ own pursuits with those of others’ into the social system as a whole. Collective coordination and integration of discrete ends and action plans is a necessary condition of cooperative rational agency and cannot be achieved without the maintenance of a commonly endorsed, socially shared
understanding of and commitment to norms and values structuring and regulating expectations individuals reciprocally have of each other. The selection of means and the pursuit of particular ends by any one agent with reference to mutual promotion of social integration is what Habermas understands by “communicative action” – “action oriented to reaching understanding and agreement” (MCCA, 58. See also TCA 1, 285-6). Such interaction is differentiated from “strategic” and “instrumental” action through which agents privilege the pursuit of their own self-interested ends and goals over the goods of collective cooperation and coordination. (Habermas’s conception of such a social order is clearly influenced by Kant’s ideal of a “Republic of Ends.”) Because any given social system is capable of breaking down in virtue of disagreement on regulative norms and behavioural expectations, morality serves the distinct purpose of conflict resolution through the adjudication of competing interests and claims.

Moral judgement … serves … to clarify legitimate behavioural expectations in response to interpersonal conflicts resulting from the disruption of our orderly coexistence by conflicts of interest.15

If what we seek is mutual understanding and agreement, any attempts at securing such ends, resolving conflicts, and the re-establishment of collective norms acceptable to all must take the form of discourse. As we have said, discursive engagement in moral deliberation and judgement is oriented to the construction of a norm that is jointly recognized to be equally in the interest of all – i.e., a norm or judgement that is universalizable or generalizable. The communicative significance of reasons proffered in argumentation rests in the fact that communicative action seeks coordination of means and integration of ends only through the giving and assessing of reasons agents present in making claims as to the generalizability of individuals’ needs, interests and goals. The shared communicative obligation here is a form of accountability that must be publicly displayed as a willingness to pursue one’s own interests only on condition interests are able to be integrated harmoniously with others’ pursuit of their own rational interests so understood. The “communicative” coordination of individuals’ ends and means is as such based on collective acceptance of the justifiability of validity claims to the rightness of projected ends and selected means for their attainment. Such acceptance of justifying

15
grounds must be intersubjectively recognized as acceptable. That I maintain my claim or interest is justified in light of reasons I give for its rightness is epistemically and morally insufficient on its own for communicative purposes of promoting or re-establishing social integration. Justifiability requires that others agree with my reasons and accept the rightness of my claims where this agreement is secured under conditions of symmetry and reciprocity. I, in turn, must have public evidence of such acceptance and be convinced of the cogency of others’ acceptance of the justifiability of my claims for a generalizable integration of interests to emerge. What we reciprocally require of each other is assurance of commitment to the dialogical rights and obligations identified by procedural conditions of discursive argumentation. Assurance of my epistemic and moral accountability as a member of the community is provided by an intersubjectively recognized ability and willingness on my part to justify the purported validity-quageneralizability of my stated interests to others. This is the “communicative” accountability that others expect of me and I reciprocally expect of them. In communicative contexts, it is understood that a justified claim to the moral rightness of ends and means is one that others rationally accept or could rationally accept in virtue of reasons offered and collectively assessed for coherence with others’ claims and interests. What Habermas terms “discourse” or “argumentation” is a conflict resolution procedure which comprises but a more systematic and rigorously pursued form of communicative action expressed by rational agency in the everyday contexts of the lifeworld: “Moral knowledge consists of a stock pile of convincing reasons for consensually settling conflicts of interaction that arise within the lifeworld” (TJ, 265).

What “good reasons” and “sound justification” mean for Habermas must be understood within this pragmatist conception of morality as a social institution. Given its functions, justification takes the dialogical form of “justification-as-justification-to-others.” “Others” here include individuals and parties potentially affected by any decision I make or support. They include those with whom one is actually engaged or could be engaged in maintaining and promoting common understandings and expectations for ends of social cooperation and integration. Habermas’s term for justification - “Rechtfertigung” - refers specifically to a form of argument and
accountability motivated by the end of justifying oneself to others, and constructing agreement on the rightness of norms and judgements equally in the interest of all. Consequently, the meaning and justification of a claim to moral rightness (generalizability) is not adequately or appropriately understood in semantic terms of truth conditions for propositions. For Habermas, such a conception is fated to understand justification in “monological” terms – i.e., with reference to truth conditions deemed to be satisfiable by a proposition independent of the discursive process. As Joseph Heath writes, on the monological conception, justifiability is thought as a possible *intrinsic* predicate or property of beliefs: a belief fails to be justified if it fails to possess whatever property is deemed to be criterial for its truth or rightness – i.e., correspondence to an external world, accordance with criteria of logical validity, predictability, universalizability of a maxim.16 On that conception, the possession of such a property (or not) is not a matter needing to be decided via discourse nor is the property itself viewed as requiring discursive construction. As we have seen Siegel and Rescher to contend, truth of belief or rightness of judgement remains itself epistemically independent of discourse; any product of discourse is itself to be assessed via external epistemic criteria. Heath importantly points out that the search for non- or a- dialogue necessary and sufficient criteria for the justifiability of beliefs has been traditional within analytic epistemology (CA, 200). In his criticisms of Weinstein’s views, Siegel’s reliance on a distinction between “propositional” and “doxastic” justification is in keeping with this tradition. A belief may be doxastically justified with reference to conditions and circumstances characterizing the holder of the belief – i.e., consistency with her other beliefs, limited access to evidence, etc. The distinction here differentiates between the justifiability of holding a belief from the justifiability of the belief itself. At times, we may have a right to be wrong. And yet the belief itself, its propositional content considered independently of agent-relativity, may be false and may justifiably be shown to be false on putatively objective epistemic criteria formulated independent of discursive requirements. This is a view of justification, as Siegel acknowledges, that is independent of reasoning and justifying agents. As Karl Popper puts it, this epistemology maintains the possibility of “knowledge without a knowing subject.” As we have seen, the view
upholds a strict separation between what comes to *count* or serve as a (good) reason and what genuinely *is* a (good) reason.

Heath recognizes that while Habermas eschews the monological terms of the traditional epistemological task of identifying necessary and sufficient conditions of truth and rightness in semantic terms, it would be a mistake to view Habermas’s Discourse Theory of Morality as an abandonment of the philosophical project of identifying the possibilities and limits of the justifiability of validity claims to moral rightness. Heath understands Habermas to be addressing the problem of identifying conditions of moral justification and rightness through a reconstruction of those principles and norms structuring and governing the procedure by which a moral claim not only comes to *count* as justified but actually comes to *be* justified in the only way it can (CA, 200). Habermas, this is to say, is not concerned simply with identifying norms of social interaction and justification presently enjoying social currency; the philosophical (epistemological) question concerns their rational justification and acceptability (JA, 38-9, 42, 59, 70, 160; GA, 42-3; TJ, 238, 258-9). The task of showing that discursive procedure under ideal conditions of symmetry and reciprocity bears epistemic status does not bequeath the philosophical problem of justification to empirical sociology or to a Rawlsian “overlapping consensus.” As Heath writes, for Habermas the assessment of a belief via an examination of whether the belief intrinsically meets certain criterial conditions formulated in non-discursive terms is a characteristic mark of “monological” justification:

[According to Habermas] … by conceiving of justifiability as a property of beliefs, one tacitly treats justification as an essentially “monological” process, that is, one involving only the agent’s cognitive states and the objects of representation (CA, 201. See also MCCA, 24; TCA I, 276-78).

Heath differentiates the monological conception of justification from Habermas’s dialogical orientation which comprehends justification to be internally or conceptually related to “justification to others” – a form of justification required for the social functions we have seen morality to serve. On pragmatist grounds, the justifiability of a
belief or claim is understood with reference to the securement of a rationally motivated agreement with and between others on its truth or rightness: “the property of being justified may only be explicable with reference to the process through which a belief can be justified [to others]” (CA, 203). Meeting epistemic criteria of moral judgement, this is to say, is not possible for the solitary thinker on her own given the social functions we have seen moral claims to serve. The Popperian ideal of “knowledge without a knowing subject” thus holds only in the sense that moral justification is necessarily an intersubjective achievement – a collective achievement not possible for any one individual subject. It is the socially grounded acceptability of a norm or claim as generalizable, acceptability governed by ideal conditions of symmetry and reciprocity that is now the source of both epistemic validity and moral authority. Habermas writes:

[The question is] whether the concept of the validity of a sentence can be explicated independently of the concept of redeeming [justifying] the validity claim raised through the utterance of the sentence. I am defending the thesis that this is not possible. Semantic investigations of … normative sentences, if only they are carried through consistently enough, force us to change the level of analysis. The very analysis of the conditions of the validity of sentences itself compels us to analyze the conditions for the intersubjective recognition of corresponding validity claims (TCA I, 316, highlighting omitted).

Accordingly, the founding principle of Habermas’s Discourse Theory of Morality reads: “(D) 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” (MCCA, 66). Principle (U) specifies what a justifiable norm or judgement means as an outcome of discourse:

(U) 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 (GA, 42).

Important to note is that contrary to both Siegel and Weinstein, procedural conditions epistemically understood are not themselves substantive moral norms; they are referred to by Habermas as “epistemic presuppositions of discourse” or “general presuppositions of argumentation” (JA, 31, 77) and sometimes as “argumentative rules of discourse” (JA,
Procedural conditions are epistemic precisely in being necessary for and universal across constructions of a generalizable norm or interest. This distinction is fundamental to the claim by strong dialogicality that moral failure in discursive argumentation and judgement is the result of procedural failure.17

III

We can get a clearer view of the distinction between monological and dialogical justification by seeing its operation in Habermas’s critique of Kant’s position that each of us, relying solely on our own individual resources, is able to judge the universalizability of a maxim as determined by application of the Categorical Imperative/Principle of Equal Respect for Persons.

There are two central aspects to the notion of the monological in Habermas’s critique of Kant’s moral theory. First, as we have already seen Heath to argue, it can refer to a specific conception of the nature of propositional belief and its justification in the semantic terms of truth conditions. But second, it refers to a particular conception of the individual’s capacities for autonomous justification of judgements, maxims and norms. Habermas’s critique identifies the operation of both aspects in Kant’s moral theory. In the former case, the idea that the justifiability of a belief hinges on the possession of a semantic property intrinsic to the belief is reflected in Kant’s view that the universalizability of a maxim constitutes precisely such a property. According to Kant, whether a maxim is universalizable is a question of whether it represents a form of willing and acting that is morally permissible (justifiable and right) or forbidden (unjustifiable and wrong). Which modality a maxim possesses is determined by investigating whether the maxim, when attempted to be universalized, itself exhibits features of self-contradiction, illegitimate self-exemption, violations of equal respect for persons, and/or fitness for universal legislation in a Kingdom of Ends.18 For Kant, as Habermas reads him, collective dialogical agreement under ideal discursive conditions on the universalizability of a maxim is not an epistemic standard. If a maxim is correctly determined to be universalizable by a rational individual agent, i.e., if it is “objectively
valid” as Kant has it, then it follows that any rational agent could accept that determination, and could will the maxim as a moral law validly applicable to all agents, agents who themselves could will the maxim as a universal law. Habermas maintains that it is an essential feature of Kant’s understanding of human capacities for moral autonomy that justification of the rightness or wrongness of a maxim can soundly be performed by any rational individual agent on his/her own. One doesn’t need to ask others for their views for purposes of corroboration of one’s own judgement. This because Kant conceives of universalizability (or not) as a property that is housed immanently within the maxim itself, and as such can be determined to be present or absent by any rational agent. We do not need others to tell us what our moral obligations are to others and ourselves. I believe Habermas is largely correct in his reading of Kant here. Agreement on the objective validity of a maxim is for Kant no doubt part of the meaning of moral right and wrong since a universalizable maxim is by definition one that all could rationally agree to will and legislate for all. But joint and mutual agreement between agents on the universalizability of a maxim under ideal conditions of symmetry and reciprocity is not the kind of “agreement” Kant has in mind. Consequently, discursive agreement can play no epistemic role within Kant’s moral theory.19

Habermas’s dialogical conception of justification rejects Kant’s assumption that each of us, as rationally autonomous individuals, bears within ourselves the necessary and sufficient epistemic resources to correctly judge the universalizability of a maxim (GA, 33). Kant’s insistence that what is morally permissible, obligatory or forbidden comprise matters that each of us can rationally and correctly identify for ourselves is internally tied to the view that the truth (or falsity) of belief or rightness (or wrongness) of a maxim is an intrinsic semantic property of the belief or maxim itself. “Objective validity” is the ground of sound agreement for Kant, not its result. Habermas denies both aspects of the monological here. He maintains that the very idea of a subject examining maxims for intrinsic universalizability - i.e., their accordance with the Categorical Imperative - reveals the monological (Cartesian) conception of subjectivity under which Kant’s moral theory labours (JA, 51). What Kant fails to recognize in offering his “egocentrically conceived universalizability test” is that no moral norm, principle or
maxim can be justified “privately in the solitary monologue of the soul with itself (JA, 7. See also 51 & 64). Both the character and locus of universalization as a test of judgement come to be mistakenly formulated once we posit the required deliberative resources to rest wholly in “the individual thinker;” “the isolated subject” (JA, 51; MCCA, 203). What is thereby occluded, Habermas argues, is that the moral status of a maxim is not identifiable as an objective, intrinsic property able to be read off by the individual rational agent reasoning alone. A maxim does not necessarily retain its self-same identity across multiple deliberating agents and culturally embedded selves, and it is a mistake to believe that we have the power to identify the property of universalizability on solely individual resources. 20

What Habermas is identifying here as the second feature of Kant’s monologism is the metaphysical view that rational human agency essentially (noumenally) possesses a metaphysical self-sameness across differencing phenomenal appearances as displayed in agents’ interests and world views. Habermas’s complaint is over the assumption “that all subjects in the Kingdom of Ends share the same conception of themselves and the world” (JA, 51). On Habermas’s diagnosis, it is on the grounds of this metaphysics that Kant was led to believe that rationality itself decrees that a verdict passed on the universalizability of a maxim by any one rational agent will and must be the same verdict any other rational agent could arrive at. As on the semantic conception, differences in the moral assessment of a maxim across different agents are attributed to failure in correctly identifying the maxim’s own intrinsic properties and moral status. Habermas resists such attribution, maintaining that we cannot take for granted that a maxim I as an individual rational thinker determine to be universalizable will (must) be found by others to possess that same status, as if others were but clones of a metaphysically identical template of rationality (JA, 7, 64). By Habermas’s lights, this metaphysical postulation “simply takes for granted a prior [unifying] understanding among a plurality of empirical egos” (MCCA, 203).

It is easy to see why the discourse principle requires this kind of practice [i.e., argumentation] for the justification of norms and value decisions: whether norms and values could find the rationally motivated assent of all those affected can be judged only
from the intersubjectively enlarged perspective of the first person plural. This perspective integrates the perspectives of each participant’s worldview and self-understanding in a manner that is neither coercive nor distorting. 21

In rejecting Kant’s monological assumptions and offering a dialogical version of the Categorical Imperative, Kant’s question of universalizability – i.e., whether I could will that all persons in relevantly similar situations ought to act in accordance with the same maxim and could themselves will the maxim as a universal law – is reformulated as the question of what all could will in common, what is in the common interest, what is equally good for all. (JA, 6-7, 24). The construction of valid (generalizable) moral norms thus requires a dialogical procedure as the emphasis now shifts from a test of universalizability from a first-person singular perspective to a first person plural perspective. Thomas McCarthy succinctly captures this dialogical reformulation of Kant’s Categorical Imperative:

Rather than ascribing as valid to all others any maxim I can will to be a universal law, I must submit my maxim to all others for purposes of discursively testing its claim to universality. 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 (McCarthy in MCCA, 67).

In Habermas’s own terms:

[T]he reflexive application of the universalization test calls for a form of deliberation in which each participant is compelled to adopt the perspectives of all others in order to examine whether a norm could be willed by all from the perspective of each person (GA, 33. See also JA, 48-49, 52; MCCA, 67; GA, 31).

Procedural conditions of symmetry and reciprocity set rigorous epistemic standards for deliberation on the generalizability of norms and interests. The construction and assessment of any candidate generalizable norm as required by (D) and in accordance with (U), requires that all participants be informed of the relevant needs, interests and value-orientations held by other participants. There can be no “good reasons” or “sound argument” independent of such information. Collective attainment of such knowledge cannot be left to individual thinkers’ capacities for imagination or reasoning ability.
actual engagement in dialogue is not simply adventitious to justifiable judgement but is epistemically required. Interpreting Habermas on this point, William Rehg writes:

[T]he relevant facts or circumstances for applying some need-related concept are to some extent unique for each individual, such that each has a (somewhat) unique perspective on those facts, a perspective no one can fully share. This uniqueness cuts both ways, of course. Consequently, accurate assessments of the force of needs-based arguments [in construction of a generalizable norm] must factor in the estimates of everyone involved. 22

The impossibility of generating a generalizable norm from an individual’s own solitary perspective also means that such norms cannot be justified via appeal to a simple additive convergence or aggregation of individually generated views. Habermas’s is not a consensus model of justification. The “general will” here is not a matter of counting heads. Given the specific functions served by the enterprise of morality, all must jointly understand that and why each of the others accepts or rejects proposed norm N’s generalizability. Each must reciprocally find intelligible the cogency of the others’ grounds for acceptance or rejection. Should I have reason to believe that an interlocutor’s understanding of say, the efficacy and/or efficiency of norm N to bring about his own or others’ stated desired ends is mistaken, or should I come to believe that some other norm P would be more fully generalizable given others’ and my own interests, the criterion of reciprocity epistemically obligates me to present my reasons for views for collective examination. This mutually held obligation is epistemic in that there is no alternative way of justifiably constructing a generalizable norm and no alternative to discursive examination for the justification and rightness (generalizability) of the norm. The obligation is symmetrical: all participants jointly bear the very same responsibility to themselves and each other. Other possible sources of disagreement and topics for examination include conflicting views on what a participant’s genuine interests or needs are given his stated long-term goals and ends. Discourse does not preclude the making of self-interested claims; nor does it deny that individuals have a certain privileged access to their own needs and interests. However, neither of these allowances entails privileged access to the generalizability of individual agents’ in solitary reflection.
[G]eneralized reciprocal perspective-taking (‘of each,’ ‘jointly by all’) requires not just empathy for, but also interpretive intervention into, the self-understanding of participants who must be willing to revise their descriptions of themselves and others (and the language in which they are formulated) (GA, 42-3; See also MCCA, 67, 104, 177, 204).

Needs and interests are subject to “need-interpretation” and examined as possible contributions to the construction of a generalizable interest. Only when all are jointly agreed that N is indeed equally in the interest of all as defined by (U), and that this judgement was arrived at under satisfactory fulfilment of procedural conditions of symmetry and reciprocity, are we justified in judging N to constitute a generalizable interest or norm. Individually proffered reasons here are essential ingredients in the procedure of argumentation even though the goodness of reasons, i.e., their epistemic warrant in the task of constructing a generalizable interest, cannot be identified and assessed independent of discursive argumentation.

Let us take a closer look at the criterion of reciprocity as it is central to the justification of strong dialogicality. For N to be generalizable norm is for N to be recognized as such. N must be jointly agreed by all to be valid for all: what is agreed upon is agreed to “by each for all and by all for each” (GA, 31. See also TJ, 268). Rehg captures well the reciprocal character of such mutual agreement:

[F]or A to consider the argument for N … [some norm or judgement] cogent she must not only be convinced for herself that this argument appropriately describes B’s interests (as well as her own); she must in addition have grounds for supposing that B himself accepts the argument in terms he finds appropriate. Conversely, B must suppose the same of A. … I must in addition have grounds to suppose this acceptance is not distorted, i.e., that it rests solely on an insight into the better argument. … To the extent that this does not hold, our norm does not rest on argued agreement and hence is not valid.”23

The epistemic criterion of reciprocity helps to explain why no single individual’s judgement on the generalizability of N can possess the epistemic warrant and moral authority required for a validly constructed generalizable norm. I believe it also explains why a collectively recognized fulfilment of procedural conditions within mutual agreement on a norm yields an epistemic superiority not available to judgements arising
solely from the perspectives of individual participants themselves. This interpretation of Habermas’s position is put forth in William Rehg’s “Better Justification Thesis.”

Even after one has engaged in discourse and reconsidered one’s views in light of others’ contributions, one cannot, as an individual, access the objective cogency of reasons as well as a group can when it has arrived at a rational consensus. According to the “Better Justification Thesis” … a reasonable consensus of reasonable individuals on some moral claim provides *ceteris paribus* a better reflection or measure, a more accurate assessment, of the objective cogency of the relevant arguments. One can, I suspect, cast the idea as a matter of degree: a judgment is better justified to the extent that reasonable individuals in discourse agree on it (GF, 122).

Rehg’s Thesis holds that it would be incorrect to view a collectively attained reason that has been generated through satisfactory fulfilment of conditions of symmetry and reciprocity to comprise substantively the same kind of reason as one generated by individual thinkers on their own. Mutual agreement, this is to say, affects “the very content of arguments;” it allows for “a better grasp of the true force of reasons” than is possible for the solitary thinker (GF, 122, 124).

[M]utual agreement adds an additional increment of justification, an addition that has the net effect of a further substantive reason, even after a discourse has put all the substantive reasons and arguments on the table (GF, 129. See also GA, 38).

Habermas clearly concurs with this reading:

In the case of a normative validity claim, the discursively achieved agreement grounds the corresponding norm’s worthiness to be recognized; to this extent the agreement itself contributes to the satisfaction of the norm’s conditions of validity.

IV

Arguments for the cogency of Habermas’s strong dialogicality require a satisfactory reply to a specifically ontological objection underwriting the semantic conception of justification. This objection reads: if the moral rightness of judgement depends upon establishment of a jointly acceptable generalizable norm or interest, then such norms or interests must somehow exist and be identifiable as such independent of
discourse in order for any epistemic assessment (and fallibility) to be possible. This ontological view countenances the view that claims to moral rightness may be justifiable despite not being submitted to discursive examination and thus not meeting the epistemic criterion as laid out by (D) and specified by (U). Moreover, it would seem to justify the view that an individual thinker may discover a generalizable interest that, while contrary to one discursively attained, is actually the true or right one. I believe this Realist view underlies the kinds of objections we have seen Siegel and Rescher to make. For on such an assumption of ontological independence, there is no reason to believe that an independent individual thinker is in principle unable to discover a generalizable interest on her own. I believe, however, that this ontologically-based objection remains recognizably monological in that it perpetuates the semantic conception of justification. For once again, we assume the possibility of justifying a norm as being generalizable via demonstration that an interest intrinsically possesses this property (or not) as an objective property external to and independent of agents deliberating and judging under discursive conditions. Habermas rightly denies the cogency of this Realist ontological objection. In response, he warns against ontologizing interests (TJ, 266-8; JA, 24). Moral rightness, Habermas argues, is itself constituted by justification as conceived under ideal procedural conditions of discourse. On this idealized conception, satisfaction of procedural conditions in the dialogical construction of the generalizability of a norm is what makes the agreed-upon norm morally justified and hence morally right. The process/product distinction subscribed to by Siegel and Rescher is here denied. In the absence of any compelling Realist conception of moral generalizability and rightness, the procedure of moral argumentation remains the sole and irreplaceable access to moral rightness itself. Rightness, as Habermas writes, is “justification-immanent” rather than “justification-transcendent” (TJ, 237-75) as understood on Realist premises. This for Habermas is the key difference between moral rightness and empirical truth. While discovery of objective reality presupposes some conceptual framework or paradigm of inquiry, truth or falsity of empirical hypotheses and propositions transcends our conceptual and linguistic horizons of inquiry. It is ultimately the world that decides truth. The moral domain, constituting a “communicative” order is not itself independent of the norms, principles, values, traditions and judgements we understand to make up that domain and our normative lives
within it. Recall our discussion above of the social functions served by morality, and the central features of intersubjective recognition and joint, mutual agreement required for harmonious integration and coordination of individuals’ interests and ends. Claims to moral rightness are “justification-immanent” in that the moral aim of social cohesion under universal egalitarianism cannot be attained and sustained through appeal to an external order of already internally constituted moral facts or to subject-independent reasons. In the case of our “post-metaphysical” understanding of moral rightness claims and the impartiality they intend, Habermas maintains that there is no epistemic criterion other than, or independent of, agreement under ideal discursive conditions that a judgement must meet in order to justifiably claim rightness and to be right. The meaning of moral rightness “consists entirely in ideal warranted assertibility;” moral rightness is exhausted by the notion of ideal warranted assertibility” (TJ, 255. See also TJ, 270-73).

Pace “weak dialogicality,” deliberation under dialogical conditions seeking joint and mutual agreement is not a sometime reliable indicator of moral rightness where this rightness is itself a matter of independent establishment of the substantive justifiability of a claim to generalizability. Rather, ideal discursive agreement is what makes a norm morally right, an interest or norm generalizable.

I see no difficulty in holding the view that ideal discursive agreement as laid out by (D) and (U) is a cogent epistemic criterion, and indeed comprises the only form in which post-metaphysical moral justification is possible, so long as we recognize the distinction between de facto agreement and rational agreement attained under satisfaction of procedural conditions. With reference to what we have been calling “weak dialogicality,” Habermas writes:

Rational discourse can be characterized as the sole access to moral insights – rather than merely as the ‘most promising’ or ‘most reliable’ – without elevating the de facto prevailing agreement in any given case to a criterion of the truth or rightness of moral judgements. Interpreted intersubjectively, ‘rational acceptability,’ if it is made to depend on an ideal procedure, is not tantamount to intersubjectively reached acceptance (TJ, 318).

The “goodness” of reasons thus remains a central normative concern:
'Rightness’ means rational acceptability supported by good reasons. … [But] substantial reasons can never ‘compel’ in the sense of logical inference or conclusive evidence. The former does not suffice for justification, because it merely explicates the content of the premises, whereas the latter is not available except in the case of singular perceptual judgments, and even then it is not beyond question (BFN, 226)

The privileging of truth conditions in justification by the semantic view cannot but leave an epistemic gap either between our premises and conclusions, or between the principles we presuppose and rely on to show the relevance of our reasons and the required justification of those principles themselves. For Habermas, our attempts to narrow such gaps can only take the form of increasing the scope of rational agreement on the generalizability of the norms and principles to which we are committed. Within these efforts, discourse constitutes a unique, irreplaceable and inescapable procedure for the impartial construction of norms equally in the interest of all. Habermas’s anti-realist moral epistemology bears significant educational force in its recognition that when our judgements are contested, our judgements resisted, it is not in virtue of their having confronted an objectively independent world. Rather is it the opposition encountered in a recognition of the limitations to the scope of that normative agreement and mutual understanding presupposed by any communicatively functioning social order.

[M]oral beliefs do not falter against the resistance of an objective world that all participants suppose to be one and the same. Rather they falter against the irresolubility of normative dissensus among opposing parties in a shared social world (TJ, 256).

The overcoming of normative dissensus through the promotion of the moral point of view as embedded in universal egalitarianism calls for an education of dispositions of thought and conduct oriented towards extending the borders of our social community and sets of accepted norms and values to include others who at present resist the given consensus (TJ, 256-7). Abilities and dispositions for seeking mutual agreement with others, a willingness to engage in the task of securing an ever-more extensive inclusion of the interests and world-views of others, can only be conceived and fostered through the cognitive, affective and conative resources discourse presupposes.
Whereas successful learning in the sphere of empirical problems may result in agreement, learning in the moral domain is assessed in terms of how inclusive such a consensus reached through reason-giving is (TJ, 257 emph. deleted).

Conclusion: Educational Implications

In this paper, I have attempted to explicate Habermas’s “strong dialogicality” as a moral epistemology in response to criticisms and doubts regarding its cogency, and I have illustrated its operation in Habermas’s critique of the limitations of certain aspects of Kant’s moral theory. I want to conclude with a few educational comments. As a moral epistemology and social theory, Habermas’s Discourse Theory of Morality, cannot on its own provide any conclusive account of educational aims and interventions. How to develop individuals’ commitments to universal egalitarianism, prescriptions of pedagogical strategies for fostering dialogical abilities, and dispositions – these matters are not the purview of philosophy and it is not as such that philosophy contributes to the educational enterprise. What a philosophical theory such as Habermas’s can do, however, is identify a framework of cognitive, conative and affective traits and dispositions of character argued to be necessary for competent engagement in discursive argumentation. Again, such an identification, grounded in philosophical argument, cannot on its own fully justify the educational value of such engagement. Only a substantive conception of the good/authentic life, the justified state, can provide such justification.

If Discourse Theory is to be applied to education, then the ideals of strong dialogicality and universal egalitarianism may serve in the work of identifying competencies, virtues and dispositions within the formulation of learning outcomes for programmes of moral and political education in pluralist democracies. Consider in this light the following amplified version of the epistemic procedural conditions, together with the rights and obligations they set for dialogical interlocutors. As these are not substantive moral rights and obligations (see note 17), my reconstruction of learning outcomes from Discourse Theory will focus only on intellectual or cognitively-structured competences and virtues required for competent participation in discourse.27
1. Equal rights to participation: publicity and inclusiveness

Practical discourses are public in nature. Nobody who could make a relevant contribution to the examination of a validity claim may be excluded. All competent speakers and actors actually or potentially affected by the resulting agreement and decision are permitted to take part in discourse (GA, 44; TJ, 106, 269). The giving and evaluation of pragmatic and ethical reasons prevents the marginalization of cultural self-understandings and worldviews held by particular individuals and groups. Discourse as such fosters a hermeneutic sensitivity to a sufficiently broad spectrum of contributions (GA, 42). This requires that discourse abide by the condition of publicity (TJ, 106-7)

2. Equal argumentational rights in participation

Everyone possesses the same opportunity to address matters under consideration; all participants are granted an equal opportunity to make contributions. Everyone is allowed to question any assertion whatever, to introduce any assertion whatever into the discourse and to express his [or her] attitudes, desires and needs. The equal communicative rights of all participants ensures that only reasons that give equal weight to the interests and evaluative orientations of everybody can influence the outcome of practical discourses (TJ, 107, 269; GA, 44).

3. Exclusion of deception and bias

Participants mean what they say and presentations and evaluations of arguments must be sincere. Participants display sincerity and honesty with oneself in the willingness to distance oneself from one’s own situation and interests in order to be able to critique self-delusions and self-deception in others’ and one’s own views. Participants must be critically alert to self-deception as well as hermeneutically open and sensitive to how others understand themselves and the world in order to provide unbiased evaluations of all arguments presented (GA, 44; TJ, 107, 269). Impartiality requires the satisfaction of the condition of reciprocity: “[e]veryone ought to put herself into everyone else’s situation and take their understanding of themselves and of the world just as seriously as her own” (TJ, 270). Participants must be willing to revise their descriptions of themselves and others (GA, 42).

4. Coercion-free communication

Communication is focussed only on epistemic factors; deliberations must be free of any restrictions that prevent the better argument from being raised. Communication must be free from external and internal coercion so that the “yes” or “no” stances that participants adopt on criticizable validity claims are motivated solely by the rational force of the better reasons. The absence of coercion and deception ensures that nothing but reasons...
can tip the balance in favour of the acceptance of a controversial norm or judgment (TJ107; GA44).

Satisfaction of these conditions within discourse requires competence both at the logical level of arguments as products and at the communicative/dialogical level of procedures ideally ensuring symmetry and reciprocity in the construction of a generalizable norm or interest. While for Habermas this distinction between product and process involves “a separation [that] … cannot be maintained” in any communicative practice (TC 1, 26), as an analytic distinction it remains useful for delineating and classifying abilities and dispositions requisite for competent engagement in discursive argumentation.

The logical level identifies and requires abilities to recognize and produce arguments satisfying formal features of logical validity, such as avoidance of contradiction and other logically fallacious forms of reasoning. Habermas would include here the recognition of informal fallacies in argument such as begging the question, equivocation, appeal to authority, etc. (MCCA 87-88). Also required is an understanding of conceptual cogency and consistency, as in: an application of predicate F to object A implies application of F to all other objects relevantly similar to A (ibid).

At the procedural level, epistemic conditions require abilities and dispositions in the areas of perspective-taking, self-knowledge, solidarity and its attitudes, and social knowledge and understanding. In the area of perspective-taking, interlocutors need to possess the ability to foresee possible social, economic and political consequences of policies being considered, together with their possible effects on specific individuals and groups if implemented. This requires a hermeneutic sensitivity to and knowledge of the cultural traditions of one’s own biography as well as those shaping the identities of others from other ethical-political traditions. In the area of self-knowledge, interlocutors display the ability to differentiate between one’s genuine long-term interests and short-term satisfactions of immediate wants and desires, and be willing to consider others’ differing views of their long-term interests. While each individual “is the last court of appeal for judging what is in his best interests” (MCCA, 67), her interests presented as claims for incorporation into a generalizable interest remain fallible and open for collective
consideration. In the area of social knowledge and understanding, individuals need to be apprised of the major social, political, ethical and moral problems and issues being debated nationally and internationally, and develop a comprehensive understanding of the basic principles, concepts, and values deployed in positions taken and arguments presented. In the area of solidarity and its attitudes, interlocutors display solidarity with others as expressed in sincere concern for the well-being and good of all interlocutors and parties potentially affected by discursive decisions. Solidarity is also displayed in the mutual interest to identify sources of disagreement and the will to find common ground in the construction of a generalizable norm. This requires a disposition of impartiality (“a decentered understanding of the world”) in which the instrumental and strategic promotion of one’s own ends is bracketed and the communicative attitude is taken up for the sake of attaining mutual and joint agreement on the generalizability of a norm or interest. Impartiality is also equally needed in the ability and willingness to articulate views and interests of interlocutors one believes to be unable to present these in a clear and comprehensive manner due to deficits in discursive competencies, and this despite disagreement with her views. An attendant virtue is self-control in refusing to engage in coercion or compulsion through appeal to one’s position of power or authority outside the arena of discourse.

These are, in broad outline, some of the intellectual skills, dispositions and commitments that marveserve as learning outcomes in a programme of moral and political education grounded in Habermas’s Discourse Theory of Morality. If my account of the cogency of this Theory and defense against objections are successful, it may serve as a partial justification of aims for that educational enterprise.

Endnotes and References


9. Kilby’s argument is particularly relevant to intercultural moral disagreement where the conflict appears to be between incommensurable conceptions of moral justification. Siegel is optimistic that such conflict can be rationally resolved given the universal applicability of epistemic criteria such as soundness. According to Siegel, the probative force of an argument … “is what it is, independently of the culture of either the arguer or her audience.” See “Multiculturalism and the possibility of transcultural educational and philosophical ideals,” *Philosophy*, #74 (1999): 387-409. 401.


15. Jurgen Habermas, Justice and application, translated by C. Cronin (Cambridge, Mass: The MIT Press, 1993) 9. Henceforth JA. See also GA, 4 & TJ, 265, 271. Since morality is a realm in which individuals interact on the basis of conceptions of selfhood, identities are also interpersonally originating and communicatively and sustained: “No one can construct an identity independently of the identifications that others make of him. These are, naturally, identifications that others make not in the propositional attitude of observers, but in the performative attitude of participants in interaction. Indeed the ego does not accomplish its self-identifications in a propositional attitude. It presents itself to itself as a practical ego in the performance of communicative actions; and in communicative action the participants must reciprocally suppose that the distinguishing-oneself-from-others is recognized by those others. Thus the basis for the assertion of one’s own identity is not really self-identification, but intersubjectively recognized self-identification.” Communication and the evolution of society, translated and with introduction by Thomas McCarthy (Boston: Beacon Press, 1979) 107.


18. Differing interpretations of the features (contradictions) Kant believed to be possessed by morally impermissible maxims are provided by Christine M. Korsgaard, “Kant’s formula of universal law,” in Creating the kingdom of ends (Cambridge: Cambridge University Press, 1996) 77-105.

19. As my purpose in examining Habermas’s critique of Kant’s moral theory is primarily to articulate the distinction between the monological and the dialogical in its relevance for moral justification, I will not here engage the hermeneutic question of the correctness of Habermas’s reading of Kant’s moral theory in defending what I consider to be an accurate reading of Kant. But see note 20 below. Habermas’s reading is contested by a number of writers. See, for example, Allen W. Wood, Kant’s ethical theory (Cambridge: Cambridge University Press, 1999): 301-3. And see Immanuel Kant, “What is orientation in thinking?” in Kant’s political writings, translated by H.B. Nisbet, edited with introduction and notes by H.S. Reis (NY & Cambridge: Cambridge University Press, 2002) 247. No doubt, there are passages in which Kant comes close to accepting strong dialogicality but at the end of the day his considered view is one of weak dialogicality.
20. I would offer the following qualification, however, as a friendly amendment I believe Habermas would accept. Kant’s monological view of maxim testing remains a correct view in the case of non-universalizable maxims displaying self-contradiction and free-riding. These violations of “strict duty” do not require engagement in discourse for a correct analysis. This qualification of course does not contradict Habermas’s view that even maxims not displaying self-contradiction and free-riding may still not be universalizable. Determination of universalizability in such cases does require discursive agreement under ideal conditions of discourse.


24. I believe Rehg’s model diverges at times from Habermas’s strong dialogicality on a number of points. It allows for the possibility that an individual’s moral judgement may be correct or right independent of actual engagement in discourse and regardless of whether it contradicts discursively attained agreement. Rehg emphasizes, however, that an individual judgement, even if correct, cannot legitimately claim fuller or more adequate *justification* for its correctness than is attained or attainable in discursively generated mutual agreement and collective judgement. But it remains the case for Rehg that the rightness of judgement is distinct from the justifiability of judgement (GF, 122). “Thus even the lone dissenter could make a substantively correct moral judgment, though that judgment will be at best only partially justified and may even be rationally defective at the time the dissenter makes it …” (GF, 123). I believe Habermas denies this distinction in his claim that moral rightness is justification-immanent, not justification transcendent.


I am grateful to the following readers for their valuable critical feedback on drafts of this paper. They may not support the views and arguments I have arrived at. James Bradley, Maeve Cooke, Andrea English, Scott Johnston, Darron Kelly, Chris Martin, Dieter Misgeld, Harvey Siegel, Evan Simpson. I am also grateful to two anonymous reviewers for Educational Theory for their suggestions.