Communicative Ethics

It is no accident that the categorical imperative is directed to the second person singular and that it creates the impression that each individual could undertake the required test of norms for himself *in foro interno*. But in fact the reflexive application of the universalisation test calls for a form of deliberation in which each participant is compelled to adopt the perspective of all others in order to examine whether a norm could be willed by all *from the perspective of each person*. This is the situation of a *rational discourse* oriented to reaching understanding in which all those concerned participate. This idea of a discursively produced understanding also imposes a greater burden of justification on the isolated judging subject than would a monologically applied universalisation test.

Kant may have been so readily inclined to foreshorten an intersubjective concept of autonomy in an individualistic direction because he failed to distinguish ethical questions sufficiently from pragmatic questions. Anyone who takes seriously questions of ethical self-understanding runs up against the stubborn cultural meaning of an individual's or a group's historically changing interpretations of the world and of themselves. As a child of the eighteenth century, Kant still thinks in an unhistorical way and consequently overlooks the layer of traditions in which identities are formed. He tacitly assumes that in making moral judgments each individual can project himself into the situation of everyone else *through his own imagination*. But when the participants can no longer rely on a transcendental preunderstanding grounded in more or less homogeneous conditions of life and interests, the moral point of view can only be realised under conditions of communication that ensure that *everyone* tests the acceptability of a norm, implemented in a general practice, also from the perspective of his own understanding of himself and of the world ... in this way the categorical imperative receives a discourse-theoretical interpretation in which its place is taken by the discourse principle (*D*), according to which only those norms can claim validity that could meet with the agreement of all those concerned in their capacity as participants in a practical discourse.

I began with the question of whether the cognitive content of a morality of equal respect and solidaristic responsibility for everybody can still be justified after the collapse of its religious foundation. In conclusion, I would like to examine what the intersubjectivistic interpretation of the categorical imperative can contribute to answering the question. Here we must treat two problems separately: First, we must clarify how much of the original intuitions a discourse ethics salvages in the disenchanted universe of postmetaphysical justification and in what sense one can still speak of the cognitive validity of moral judgments and positions (VIII). Second, there is the final question of whether the content of a morality that results from the rational reconstruction of traditional, religious intuitions remains bound, in spite of its procedural character, to it original context (IX)

VIII

With the devaluation of the epistemic authority of the God's eye view, moral commands lose their religious as well as their metaphysical foundation. This development also has implications for discourse ethics; it can neither defend the full moral contents of religious intuitions (1) nor can it represent the validity of moral norms in realist terms (2).

(1) The fact that moral practice is no longer tied to the individual's expectation of salvation and an exemplary conduct of life through the person of a redemptive God and the divine plan for salvation has two unwelcome consequences. On the one hand, moral knowledge becomes detached from moral motivation, and on the other, the concept of morally right action becomes differentiated from the conception of a good or godly life.

Discourse ethics correlates ethical and moral questions with different forms of argumentation, namely, with discourses of self-clarification and discourses of normative justification (and application), respectively. But it does not thereby reduce morality to equal treatment; rather, it takes account of both the aspects of justice and that of solidarity. A discursive agreement depends simultaneously on the nonsubstitutable "yes" or "no" responses of each individual and on overcoming the egocentric perspective, something that all participants are constrained to do by an argumentative practice designed to produce agreement of an epistemic kind. If the pragmatic features of discourse make possible an insightful process of opinion- and will-formation that guarantees both of these conditions, then the rationally motivated "yes" or "no" responses can take the interests of each individual into consideration without breaking the prior social bond that joins all those who are oriented toward reaching understanding in a transsubjective attitude.

However, uncoupling morality from questions of the good life leads to a motivational deficit. Because there is no profane substitute for the hope of personal salvation, we lose the strongest motive for obeying moral commands. Discourse ethics intensifies the intellectualistic separation of moral judgment from action even further by locating the moral point of view in rational discourse. There is no direct route from discursively achieved consensus to action. Certainly, moral judgments tell us what we should do, and good reasons affect our will; this is shown by the bad conscience that "plagues" us when we act against our better judgment. But the problem of weakness of will also shows that moral insight is based on the weak force of epistemic reasons and, in contrast with pragmatic reasons, does not itself constitute a rational motive. When we know what it is morally right for us to

do, we know that there are no good (epistemic) reasons to act otherwise. But that does not mean that other motives will not prevail.

With the loss of its foundation in the religious promise of salvation, the meaning of normative obligation also changes. The differentiation between strict duties and less binding values, between what is morally right and what is ethically worth striving for, already sharpens moral validity into a normativity to which impartial judgment alone is adequate. The shift in perspective from God to human beings has a further consequence. "Validity" now signifies that moral norms could win the agreement of all concerned, on the condition that they jointly examine in practical discourse whether a corresponding practice is in the equal interest of all. This agreement expresses two things: the fallible reason of *deliberating* subjects who convince one another that a hypothetically introduced norm is worthy of being recognized, and the freedom of *legislating* subjects who understand themselves as the authors of the norms to which they subject themselves as addressees. The mode of validity of moral norms now bears the traces both of the fallibility of the discovering mind and of the creativity of the constructing mind.

(2) The problem of in which sense moral judgments and attitudes can claim validity reveals another aspect when we reflect on the essentialist statements through which moral commands were previously justified in a metaphysical fashion as elements of a rationally ordered world. As long as the cognitive content of morality could be expressed in assertoric statements, moral judgments could be viewed as true or false. But if moral realism can no longer be defended by appealing to a creationist metaphysics and to natural law (or their surrogates), the validity of moral statements can no longer be assimilated to the truth of assertoric statements. The latter state how things are in the world; the former state what we should do.

If one assumes that, in general, sentences can be valid only in the sense of being "true" or "false" and further that "truth" is to be understood as correspondence between sentences and facts, then every validity claim that is raised for a nondescriptive sentence necessarily appears problematic. In fact, modern moral scepticism is based on the thesis that normative statements cannot be true or false, and hence cannot be justified, because there is no moral order, no such things as moral objects or facts. On this received account, the concept of *the world* as the totality of facts is connected with a correspondence notion of *truth* and a semantic conception of *justification*. I will very briefly discuss these questionable premises in reverse order.

A sentence or proposition is justified on the semantic conception if it can be derived from basic sentences according to valid rules of inference, where a class of basic sentences is distinguished by specific (logical, epistemological, or psychological) criteria. But the foundationalist assumption that there exists such a class of basic sentences whose truth is immediately accessible to perception or to intuition has not withstood linguistic arguments for the holistic character of language and interpretation: every justification must at least *proceed* from a pre-understood context or background understanding. This failure of foundationalism recommends a pragmatic conception of justification as a public practice in which criticizable validity claims can be defended with good reasons. Of course, the criteria of rationality that determine which reasons count as good reasons can themselves be made a matter for discussion. Hence procedural characteristics of the process of argumentation itself must ultimately bear the burden of explaining why results achieved in a procedurally correct manner enjoy the presumption of validity. For example, the communicative structure of rational discourse can ensure that all relevant contributions are heard and that the unforced force of the better argument alone determines the "yes" or "no" responses of the participants.

The pragmatic conception of justification opens the way for an epistemic concept of truth that overcomes the well-known problems with the correspondence theory. The truth predicate refers to the language game of justification, that is, to the public redemption of validity claims. On the other hand, truth cannot be identified with justifiability or warranted assertability. The "cautionary" use of the truth predicate — regardless of how well "*p*" is justified, it still may not be true — highlights the difference in meaning between "truth" as an irreducible property of statements and "rational acceptability" as a context-dependent property of utterances. This difference can be understood within the horizon of possible justifications in terms of the distinction between "justified in our context" and "justified in every context." This difference can be cashed out in turn through a weak idealization of our processes of argumentation, understood as capable of being extended indefinitely over time. When we assert "*p*" and thereby claim truth for "*p*" we accept the obligation to defend "*p*" in argumentation — in full awareness of its fallibility — against all future objections.

In the present context I am less interested in the complex relation between truth and justification than in the possibility of conceiving truth, purified of all connotations of correspondence, as a special case of validity, where this *general* concept of validity is introduced in connection with the discursive redemption of validity claims. In this way we open up a conceptual space in which the concept of normative, and in particular moral, validity can be situated. The rightness of moral norms (or of general normative statements) and of particular normative injunctions based on them can then be understood as analogous to the truth of descriptive statements. What unites these two concepts of validity is the procedure of discursively redeeming the corresponding validity claims. What separates them is the fact that they refer, respectively, to the social and the objective worlds.

The social world, as the totality of legitimately ordered interpersonal relations, is accessible only from the *participant's* perspective; it is intrinsically historical and hence has, if you will, an ontological constitution different from that of the objective world which can be described from the *observer's* perspective. The social world is inextricably interwoven with the intentions and beliefs, the practices and languages of its members. This holds in a similar way for *descriptions* of the objective world but not for this world itself. Hence the discursive redemption of truth claims has a different meaning from that of moral validity claims: in the former case, discursive agreement *signifies* that the truth conditions of an assertoric proposition, interpreted in terms of assertability conditions, are fulfilled; in the latter case, discursive agreement *justifies* the claim that a norm is worthy of recognition and thereby itself contributes to the fulfillment of its conditions of validity. Whereas rational acceptability merely *points to* the truth of assertoric propositions, it makes a *constructive* contribution to the validity of moral norms. The moments of construction and discovery are interwoven in moral insight differently than they are in theoretical knowledge.

What is not at our disposal here is the moral point of view that imposes itself upon us, not an objective moral order assumed to exist independently of our descriptions. It is not the social world as such that is not at our disposal but the structure and procedure of a process of argumentation that facilitates both the production and the discovery of the norms of well-ordered interpersonal relations.

The constructivist meaning of moral judgments, understood on the model of self-legislation, must not be forgotten; but it must not obliterate the epistemic meaning of moral justifications either.

IX

Discourse ethics defends a morality of equal respect and solidaristic responsibility for everybody. But it does this in the first instance through a rational reconstruction of the contents of a moral tradition whose religious foundations have been undermined. If the discourse-theoretical interpretation of the categorical imperative remained bound to the tradition in which it originates, this genealogy would represent an obstacle to the goal of demonstrating the cognitive content of moral judgments *as such.* Thus it remains to provide a theoretical justification of the moral point of view itself.

The discourse principle provides an answer to the predicament in which the members of *any* moral community find themselves when, in making the transition to a modern, pluralistic society, they find themselves faced with the dilemma that though they still argue with reasons about moral judgments and beliefs, their substantive background consensus on the underlying moral norms has been shattered. They find themselves embroiled in global and domestic practical conflicts in need of regulation that they continue to regard as moral, and hence as rationally resolvable, conflicts; but their shared ethos has disintegrated. The following scenario does not depict an "original position" but an ideal-typical development that could have taken place under real conditions.

I proceed on the assumption that the participants do not wish to resolve their conflicts through violence, or even compromise, but through communication. Thus their initial impulse is to engage in deliberation and work out a shared *ethical* self-understanding on a secular basis. But given the differentiated forms of life characteristic of pluralistic societies, such an effort is doomed to failure. The participants will soon realize that the critical appropriation of their strong evaluations leads to competing conceptions of the good. Let us assume that they nevertheless remain resolved to engage in deliberation and not to fall back on a mere *modus vivendi* as a substitute for the threatened moral way of life.

In the absence of a substantive agreement on particular norms, the participants must now rely on the "neutral" fact that each of them participates in *some* communicative form of life which is structured by linguistically mediated understanding. Since communicative processes and forms of life have certain structural features in common, they could ask themselves whether these features harbor normative contents that could provide a basis for shared orientations. Taking this as a clue, theories in the tradition of Hegel, Humboldt, and G. H. Mead have shown that communicative actions involve shared presuppositions and that communicative forms of life are interwoven with relations of reciprocal recognition, and to this extent, both have a normative content. These analyses demonstrate that morality derives a genuine meaning, independent of the various conceptions of the good, from the form and perspectival structure of unimpaired, intersubjective socialization.

To be sure, structural features of communicative forms of life alone are not sufficient to justify the claim that members of a particular historical community ought to transcend their particularistic value-orientations and make the transition to the fully symmetrical and inclusive relations of an egalitarian universalism. On the other hand, a universalistic conception that wants to avoid false abstractions must draw on insights from the theory of communication. From the fact that persons can only be individuated through socialization it follows that moral concern is owed equally to persons both as irreplaceable individuals and as members of the community, and hence it connects justice with solidarity. Equal treatment means equal treatment of unequals who are nonetheless aware of their interdependence. Moral universalism must not take into account the aspect of equality — the fact that persons as such are equal to all other persons — at *the expense of* the aspect of individuality — the fact that as individuals they are at the same time absolutely different from all others. The equal respect for everyone else demanded by a moral universalism sensitive to difference thus takes the form of a *nonleveling* and *nonappropriating inclusion* of the other in his *otherness.*

But how can the transition to a posttraditional morality as such be justified? Traditionally established obligations rooted in communicative action do not *of themselves* reach beyond the limits of the family, the tribe, the city, or the nation. However, the reflexive form of communicative action behaves differently: argumentation of its very nature points beyond all particular forms of life. For in the pragmatic presuppositions of rational discourse or deliberation the normative content of the implicit assumptions of communicative action *is generalized, abstracted, and freed from all limits* — the practice of deliberation is extended to an inclusive community that does not in principle exclude any subject capable of speech and action who can make relevant contributions. This idea points to a way out of the modern dilemma, since the participants have lost their metaphysical guarantees and must so to speak derive their normative orientations from themselves alone. As we have seen, the participants can only draw on those features of a common practice they already currently share. Given the failure to identify a shared good, such features shrink to the fund of formal features of the performatively shared situation of deliberation. The bottom line is that the participants have all already entered into the cooperative enterprise of rational discourse.

Although it is a rather meager basis for justification, the neutral content of this common store may provide an opportunity, given the predicament posed by the pluralism of worldviews. A prospect of finding an equivalent for the traditional, substantive grounding of a normative consensus would exist if the *form of communication* in which joint practical deliberation takes place were such that it makes possible a justification of moral norms convincing to all participants because of its impartiality. The missing "transcendent good" can be replaced in an "immanent" fashion only by appeal to the intrinsic constitution of the practice of deliberation. From here, I suggest, three steps lead to a theoretical justification of the moral point of view.

(a) If the practice of deliberation itself is regarded as the only possible resource for a standpoint of impartial justification of moral questions, then the appeal to moral content must be replaced by the self-referential appeal to the form of this practice. This is precisely what is captured by:

(D) Only those norms can claim validity that could meet with the acceptance of all concerned in practical discourse.

Here the "acceptance" (*Zustimmung*) achieved under conditions of rational discourse signifies an agreement (*Einverstandnis*) motivated by epistemic reasons; it should not be understood as a contract (*Vereinbarung*) that is rationally motivated from the egocentric perspective of each participant. On the other hand, the principle of discourse leaves open the type of argumentation, and hence the route, by which a discursive agreement can be reached. (D) does not by itself state that a justification of moral norms is possible without recourse to a substantive background consensus.

(b) The hypothetically introduced principle (D) specifies the condition that valid norms would fulfill if they *could* be justified. For the moment we are only assuming that the concept of a moral norm is clear. The participants also have an intuitive understanding of how one engages in argumentation. Though they are assumed only to be familiar with the justification of descriptive sentences and not yet to know whether moral validity claims can be judged in a similar way, they can form a conception (without prejudging the issue) of what it would mean to justify a norm. But what is still needed for the operationalization of (D) is a rule of argumentation specifying how moral norms can be justified.

The principle of universalization (U) is indeed inspired by (D), but initially it is nothing more than a proposal arrived at abductively.

(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.

Three aspects of this formulation are in need of clarification. The phrase "interests and value-orientations" points to the role played by the pragmatic and ethical reasons of the individual participants in practical discourse. These inputs are designed to prevent the marginalization of the self-understanding and worldviews of particular individuals or groups and, in general, to foster a hermeneutic sensitivity to a sufficiently broad spectrum of contributions. Second, generalized 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). Finally, the goal of "uncoerced joint acceptance" specifies the respect in which the reasons presented in discourse cast off their agent-relative meaning and take on an epistemic meaning from the standpoint of symmetrical consideration.

(c) The participants themselves will perhaps be satisfied with this (or a similar) rule of argumentation as long as it proves useful and does not lead to counterintuitive results. It must turn out that a practice of justification conducted in this manner selects norms that are capable of commanding universal agreement — for example, norms expressing human rights. But from the perspective of the moral theorist there still remains one final justificatory step.

We may assume that the practice of deliberation and justification we call "argumentation" is to be found in all cultures and societies (if not in institutionalized form, then at least as an informal practice) and that there is no functionally equivalent alternative to this mode of problem solving. In view of the universality and nonsubsititutibility of the practice of argumentation, it would be difficult to dispute the neutrality of the discourse principle (D). But ethnocentric assumptions, and hence a specific conception of the good that is not shared by other cultures, may have insinuated themselves into the abduction of (U). The suspicion that the understanding of morality operationalized in (U) reflects eurocentric prejudices could be dispelled through an "immanent" defense of this account of the moral point of view, that is, by appealing to knowledge of what it means to engage in the practice of argumentation as such. Thus the discourse-ethical model of justification consists in the derivation of the basic principle (U) from the implicit content of universal presuppositions of argumentation in conjunction with the conception of normative justification in general expressed in (D).

This is easy to understand in an intuitive way (though any attempt to provide a formal justification would require involved discussions of the meaning and feasibility of "transcendental arguments"). Here I will limit myself to the observation that we engage in argumentation with the intention of convincing one another of the validity claims that proponents raise for their statements and are ready to defend against opponents. The practice of argumentation sets in motion a *cooperative* competition for the better argument, where the orientation to the goal of a communicatively reached agreement unites the participants from the outset. The assumption that the competition can lead to "rationally acceptable," hence "convincing," results is based on the rational force of arguments. Of course, what counts as a good or a bad argument can itself become a topic for discussion. Thus the rational acceptability of a statement ultimately rests on reasons in conjunction with specific features of the process of argumentation itself. The four most important features are: **(i)** that nobody who could make a relevant contribution may be excluded; **(ii)** that all participants are granted an equal opportunity to make contributions; **(iii)** that the participants must mean what they say; and **(iv)** that communication must be freed 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. If everyone who engages in argumentation must make at least these pragmatic presuppositions, then in virtue of (i) the public character of practical discourses and the inclusion of all concerned and (ii) the equal communicative rights of all participants, only reasons that give equal weight to the interests and evaluative orientations of everybody can influence the outcome of practical discourses; and because of the absence of (iii) deception and (iv) coercion, nothing but reasons can tip the balance in favor of the acceptance of a controversial norm. Finally, on the assumption that participants reciprocally impute an orientation to communicative agreement to one another, this "uncoerced" acceptance can only occur "jointly" or collectively.

Against the frequently raised objection that this justification is circular I would note that the content of the universal presuppositions of argumentation is by no means "normative" in the moral sense. For inclusivity only signifies that access to discourse is unrestricted; it does not imply the universality of binding norms of action. The equal distribution of communicative freedoms and the requirement of truthfulness *in discourse* have the status of *argumentative* duties and rights, not of *moral* duties and rights. So too, the absence of coercion refers to the process of argumentation itself, not to interpersonal relations *outside* of this practice. These constitutive rules of the language game of argumentation govern the exchange of

arguments and of "yes" or "no" responses; they have the epistemic force of enabling conditions for the justification of statements but do not have any *immediate* practical effects in motivating actions and interactions outside of discourse.

The point of such a justification of the moral point of view is that the normative content of this epistemic language game is transmitted only by a rule of argumentation to the selection of norms of action, which together with their moral validity claim provide the input into practical discourses. A moral obligation cannot follow from the so to speak transcendental constraint of unavoidable presuppositions of argumentation alone; rather it attaches to the specific objects of practical discourse, namely, to the norms *introduced* into discourse to which the reasons mobilized in deliberation refer. I emphasize this when I specify that (U) can be rendered plausible *in connection with a* (weak, hence nonprejudicial) *concept of normative justification.*

This justification strategy, which I have here merely sketched, must be supplemented with genealogical arguments drawing on premises of modernization theory, if (U) is to be rendered plausible. With (U) we reassure ourselves in a reflexive manner of a residual normative substance which is preserved in posttraditional societies by the formal features of argumentation and action oriented to reaching a shared understanding. This is also shown by the procedure of establishing universal presuppositions of argumentation by demonstrating performative self-contradictions, which I cannot go into here.

The question of the application of norms arises as an additional problem. The principle of appropriateness developed by Hans Günther first brings the moral point of view to bear on singular moral judgments in a *complete* manner. The outcome of successful discourses of justification and application *shows* that practical questions are differentiated by the sharply defined moral point of view; moral questions of well-ordered interpersonal relations are separated from pragmatic questions of rational choice, on the one hand, and from ethical questions of the good or not misspent life on the other. It has become clear to me in retrospect that (U) only operationalized a more comprehensive principle of discourse with reference to a particular subject matter, namely, morality. The principle of discourse can also be operationalized for other kinds of questions, for example, for the deliberations of political legislators or for legal discourses.

Further Reading: Theory of Knowledge as Social Theory | Biography | Ethics

Jürgen Habermas Archive
Philosophy Archive @ marxists.org