

Commentary on Thomas Pogge's "On O'Neill on Kant"

by Christine M. Korsgaard

Thomas Pogge has focused our attention on two claims that are central to Onora O'Neill's version of Kantian constructivism. The first is her claim that we should regard obligations as more fundamental than rights, and the second is her specific proposal about how moral principles are to be constructed - or, to put it the old fashioned way, how the universalizability test is supposed to work. In each case Pogge has criticisms to offer and alternatives to suggest. Since O'Neill is here to defend herself I will not make that my main task. Instead I will address the alternatives Pogge has proposed, with a view to determining whether they retain what is most valuable in the Kantian ethical system.

I. Obligations and Rights

Pogge devotes most of his paper to an argument which is intended to show that there is no reason to regard duties or obligations as more fundamental than rights. Instead these are two interchangeable idioms for describing the same thing. There are two main parts to his argument. First, he distinguishes three things which one might mean by the claim that duty or obligation is fundamental, and identifies the one he takes O'Neill to have in mind. Second, he attempts to show that in this sense we can just as well say that rights are fundamental, since for every duty we can identify or construct a correlative right. I will examine each part of his argument in turn.

I begin with the three senses in which duty might be thought fundamental. First, according to Pogge, one might mean that duty is "what ultimately matters morally." Pogge takes this to be equivalent to the claim that duty is the one thing of unconditional value, and he notes that Kant gives this role not to duty but to the good will. In my view there are problems with what Pogge says about this issue. Kant himself does not equate the view that the good will is the only thing of unconditional value with the view that it is "what ultimately matters morally." He does equate it, as Pogge's own

more careful formulations suggest, with the view that the good will is the condition of all value. Pogge believes that even this implies a number of distressing conclusions, such as that the happiness of innocent children is without value, and that the sufferings of evil people have positive value or at least are not bad. I don't think that either of these conclusions follows from the view that the good will is the condition of all value, or, as I will call it, the conditionality thesis. For one thing, in the *Metaphysical Principles of Virtue* Kant forbids us ever to completely deny another human being all moral worth (463-464), and this means that not only children but even those who do evil things are not only potentially but also presumptively good. Pogge supposes that there is a problem if the child dies young because she will never display virtue proportionate to her happiness. Actually, however, the conditionality thesis is separable from the idea that happiness should be *proportional* to virtue. It is hard to make out from the text where the idea of proportionality even comes from, and there is room for doubt whether either notion is quantifiable in the way that proportionality requires. Even if the conditionality thesis does call for a proportionality between happiness and virtue, however, there is no reason to suppose that the amount of virtue one has bears any special relationship to the amount of virtue one has the opportunity to display. Now admittedly these considerations do not completely lay Pogge's worry to rest, for it seems to many of us as if there is value in the happiness, or at all events the cheerfulness, contentment, security, and playfulness, of non-human animals, and the happiness of children seems to have a value of exactly this kind. Relatedly we may feel that the pains and terrors of innocent creatures are bad in a rather special way. I think that these facts raise complex issues for Kant's theory of value, although I do not share Pogge's conviction that it cannot accommodate them. It would be out of place to pursue these problems here, but I do want to say something about the supposed implication that it is good or at least not bad if evil people suffer. So let us turn briefly to Pogge's case of the sufferings of the youthful Hitler. A first and very important point is that it simply does not follow from a theory which ties the value of happiness to a person's goodness that the *sufferings* of evil people are *positively good*. Nothing follows about the sufferings of evil people: all that follows is that

their happiness is worthless. The familiar view that it is appropriate for evil people to suffer goes with a different idea, the idea of desert, which I believe cannot be explained in terms of value theory alone. As things stand, our intuitions about whether the sufferings of the evil are good, bad, or indifferent are rather tangled up with this other notion, and a proper examination of the question would require that we analyze the notion of desert and separate its implications for value theory from other ideas. Unfortunately desert is one of the few moral notions that Kant himself uses without analysis and thoughts about desert may have played an uncontrolled role in his own presentation of his value theory. The conditionality thesis is best understood as a practical principle: if one's own happiness is pursued in a way that is not consistent with the value of other human beings, then it is worthless. The way in which the will is exercised conditions the value of its objects. Kant's talk of a spectator not being pleased by the sight of the prosperity of the wicked may therefore be misleading. There is much more to say, but I must not linger over issues of value theory here, for Pogge correctly thinks that O'Neill is not assigning obligation either the role of the unconditional source of value or the role of what ultimately matters morally.

The second sense of priority Pogge identifies concerns the content of moral concepts: the most fundamental moral concept is the one in terms of which the others are defined. On this interpretation O'Neill is saying that duty or obligation is the prior notion in the sense that rights, the good will, and so on, are all defined in terms of it: the good will is defined as a will committed to acting from duty, rights would be defined as claims arising from certain duties, and so on. This is the sense of priority Pogge believes O'Neill actually has in mind when she claims that obligation, or, as he would have it, duty, is more fundamental.

The third sense of priority Pogge identifies concerns justification. In foundationalist theories, he says, this kind of priority normally coincides with the second kind: the basic concept in terms of which others are defined is also the one that justifies our belief that morality is real or that grounds the existence of morality in the world. According to Pogge, foundationalist theories consist of self-evident

principles together with their logical implications. I am not familiar with any ethical theory which takes this *Principia Mathematica*-like form. Dogmatic Rationalists like Samuel Clarke and W.D. Ross did believe in self-evident moral principles, but their “theories” consisted mainly in announcing that such things exist and then defending that claim against critics, not in deriving implications from the principles. With the important exceptions of Richard Price, and in a way Christian Wolff, the moral philosophers who have invested a lot in claims about self-evidence have always been mainly polemical writers. But there *are* theories which fit Pogge’s more general description of foundationalism. The dogmatic rationalists of the eighteenth century defined the other moral concepts in terms of obligation, and they did also treat obligation as an irreducible and unassailable moral fact about the world. And the traditional sentimentalists who defined other moral concepts in terms of approval, similarly thought that the feeling of approval was the basic fact about us from which the moral world evolved. Pogge supposes that O’Neill cannot be giving priority to obligation in this foundationalist sense because he supposes that constructivism essentially involves the method of reflective equilibrium and is so is essentially anti-foundationalist. (He does not use the term “reflective equilibrium,” but it seems to be what he is describing.) O’Neill herself describes constructivism as anti-foundationalist because it relies on plausible assumptions about rationality and agency and because “the parts are ... put together with an eye to the coherence and functioning of the structure.” (CR 194) I do not believe that constructivism requires reflective equilibrium or that it is necessarily anti-foundationalist. As I understand constructivism, its distinguishing feature - and its Kantian legacy - is the view of ethics as a discipline of practice, rather than as a kind of applied theory. Ethical concepts and principles are defined functionally, in terms of the problems whose solutions they provide. In Rawls’s version, for instance, principles of justice are principles for solving the problem of how to distribute social goods. In Kant’s ethics, principles of morality are principles for solving the problem of what an autonomous being can treat as a sufficient reason for action. In O’Neill’s interpretation of Kant, principles of reason generally are principles which can govern the shared deliberations of a plurality of thinkers or agents.

A hypothetical situation in which the problem is faced in its purity may be employed in the search for a solution: Rawls's original position, or Kant's vision of a free will confronted by the choice between morality and self-love, are examples of this. Constructivist accounts of principles so conceived are subject to two sorts of openness. First, a constructivist may be open-minded about her results in this sense: she has found a principle which solves the problem, but it is possible that a better solution could be found. Second, she may be open to the possibility that she has misdescribed the original choice situation and so in a sense the original problem. In reflective equilibrium, this possibility is supposed to be discovered when the solution - the principles chosen - do not match one's considered judgments or intuitions. Either of these two kinds of openness *may* characterize constructivism, but neither has to. One may have described the original problem in the only intelligible way, and a solution may follow inevitably from the description. Although not founded on self-evident truths, morality would then be founded on the existence of things we cannot deny. Kant thinks, for instance, that we cannot deny our autonomy: we must act under the idea of freedom. And if the categorical imperative is really the law of autonomy then it is the unique solution to the problem autonomy presents. In this way, foundations can be practical as well as theoretical. What constructivist approaches have in common is the rejection of theoretical or epistemological foundations for morality, not the rejection of foundations in general.

Although I do not have time to expand on it, this is the right moment to register a small squeak of protest against Pogge's characterization of Kant as founding morality on "extra-moral notions of freedom, agency, and rationality." For one thing, by Pogge's criteria this means that Kant is not a constructivist. I think instead that Kant is both a foundationalist and a constructivist in the sense I have just been describing. In the second place, and relatedly, there is some reason to doubt whether the notions of freedom, agency, and rationality which Kant employs *are* extra-moral. If O'Neill is correct in thinking that the categorical imperative is not only the principle of moral reasoning, but also the principle of reasoning generally, it is not even clear that the idea of an extra-moral notion of rationality makes sense. These are deep matters: according to Kant as O'Neill construes him, even the principles

of theoretical reasoning arise from our attempt to construct a world which we can share. In a sense, on this view rationality is grounded in morality rather than the reverse.

I will come back to the question whether obligation plays a foundational role in O'Neill's account. Pogge, as I have mentioned, takes O'Neill to mean that obligation has priority in content: we can define rights in terms of obligations, but not all obligations in terms of rights. It is to this view that his counter-argument is addressed. Specifically, his argument hinges on a separation of levels. Pogge believes that it is helpful to describe things this way: we have generative duties which give rise to particular obligations; and generative rights which give rise to particular claims. General duties are correlative with general rights; and particular claims are correlative with particular obligations. At the moments when O'Neill denies that correlatively always holds, she is usually denying that correlatively holds between a general duty and a particular claim, and therefore is guilty of comparing apples and oranges. Pogge argues that we *can* construct a right correlative to our duty to help the destitute. It is the right of the destitute to be helped by those who are not already doing enough. He concludes that it does not matter, except possibly for rhetorical purposes, whether we privilege the language of rights or of obligations.

A first worry concerns the formulation of this right. The claims to which it gives rise are levied against "those who are not already doing enough." This supposes that we have some clear idea how much one should be doing, and so denies one of the ways in which imperfect duties are imperfect - they do not specify exactly how much we should do. This is not a small problem if our account is to be Kantian, because this kind of imperfection results from the way in which the duty is generated from the universalizability test. One tests the maxim of not helping anyone ever, finds it to be unacceptable, and concludes that one must help someone sometime. To conclude that one has a perfect duty to give "enough", we must first specify how much is enough. Then we must show that the maxim of giving less than that amount (or, in Pogge's test, of being free to give less) cannot be universalized. It would be interesting to see whether and how this could be done.

Leaving that aside, I should like to say something about why I think that it does matter whether one talks about obligations or rights. Let me begin by making a general point about the use of moral terms like “rights” “duties” and “obligations.” I suppose that in dealing with these notions the claims of ordinary language philosophy can never have been very tempting. While the languages in which we work provide us with a wealth of moral terms, there seems to be a great deal of variation as to how exactly they are used, and they do not seem to present us with a well-defined map of the moral territory. When a moral philosopher proposes to fix these terms - say by using “obligation” only for particular duties belonging to individuals, or only for acquired duties - the proposal is usually self-consciously revisionary and pragmatic. We find ourselves in possession of a bunch of terms and in need of making a bunch of distinctions, and we decide that a certain way of lining them up will be of use. One kind of consideration in favor of using a given term in a given way is political and rhetorical: in her book on hunger, for instance, O'Neill points out some of the advantages of using the language of rights to empower the needy and lend a sense of urgency to their claims. A closely related kind of consideration is admonitory: some terms are better suited to keep us in mind of facts we are inclined to forget. O'Neill adduces a consideration of this kind when she points out that, since rights are things we envision individuals demanding for themselves, talk of children's rights may obscure the fact that children are not in a position to demand anything for themselves.

But of course there are also more straightforwardly philosophical considerations. The moral territory is complex, and the more distinctions we can capture with our terminology the better our chances will be of describing our situation correctly. One of the things I find valuable in Kant's ethics is the rather elaborate classification of different kinds of duties which he provides. Perfect and imperfect duties, duties of justice and of virtue, duties of respect and of love, have somewhat different motivational origins and different social and legal implications.

Essentially there are two benefits to such a classification. The first concerns our decisions about what to do in cases where there is apparent conflict. If the duties are of different types, and we have a

theory about which type of duty has priority and why, some guidance through conflicts will be provided. In Kant's view, for example, duties of justice have priority over the special duties of virtue; perfect duties have priority over imperfect ones; and considerations of respect normally have more weight than considerations of benevolence. No doubt even this is too crude, but it does provide a kind of guidance which we lose if all duties are supposed to be on the same plain.

The other benefit of classification has to do with the implications of an action's being right or wrong. Nothing has done more damage to the possibility of refined moral thinking than the beliefs that every action is either simply "obligatory," "permissible," or "wrong", and that the obligatory ones can always properly be demanded by others and the wrong ones always properly forbidden by law. Many people engaged in the current debate about abortion, to take one salient example, seem to think that the question whether abortion should be illegal is settled by the answering the question whether it is wrong. According to a Kantian account this is not necessarily so. Some duties give rise to enforceable claims, some to claims that are not enforceable, and some to no claims at all. Classification also helps with our own self-understanding and our understanding of one another. There are different ways of going wrong, and no doubt some people are more inclined to specific kinds of moral defect. The inclination to ride roughshod over people's rights is not the same moral defect as persistent insensitivity to vulnerability, and neither of these is the same moral defect as the raw tendency to forget about the suffering that one does not see. The more refinement we retain in our moral categories, the more different things there are to say about actions, cases, virtues, vices, that really seem to be different.

Kant himself uses the idea of a "right" in a quite specific way. To say that someone has a right is to say, roughly, that he has a claim against others which may be coercively enforced, and, by implication, which may be legally enforced by the state. Enforceable rights arise from the legitimacy of protecting one's freedom: coercion is wrong in all cases but one, and that is when it is used to fight coercion itself. It follows from this conception that rights arise, so to speak, in a specific context - the context of interaction, in which people *can* pose threats to each other's freedom. This is an important

point which I think some other conceptions of justice leave out. Kant's political thought moves from individual interaction to the nation-state and then outward, supposing that relations of justice between individuals who are far away from each other only come to be at issue when their respective nations engage in trade or war. Issues of justice arise among the reciprocally dependent, among those who do things together. Of course this does not mean that the distant can have no obligations towards each other: the fact of common humanity is sufficient to ground those. But interaction must take place before issues of *justice* can arise.

The people O'Neill addresses, whom she thinks depend too heavily on the language of rights to discuss the problems of hunger or children, are not all using "right" in so specific a sense. And Pogge denies that the ideas of right and enforceability have to go together. But there are benefits to reserving terms for enforceable claims and for the special claims which those who interact have against each other. As O'Neill has often pointed out, liberals may argue that some issues of welfare can plausibly be construed as issues of justice. People who are destitute are subject to coercion and intimidation, and this means that they are not only needy but substantively unfree. If coercion can be exercised in the name of freedom, coercion can be exercised to alleviate their condition. In this case needy people have a right, as against those with whom they interact, to the satisfaction of their basic needs. One can accept this argument, as I do, and still agree with O'Neill that quite apart from it we have an obligation to feed the hungry anywhere. But the differences still matter. The hungry in our own community, victims of unequal opportunity or an unjust economic system in which we share, have a different kind of claim on us than the hungry elsewhere do. The hungry in other nations are sometimes more directly the victims of injustice within their own communities, although too often it is injustice which our government or economic practices have supported and encouraged. Still others may not be the victims of injustice at all: they may simply be poor or the victims of catastrophe. Here our obligations are of charity, not justice. To say that the morally significant fact in all of these cases is that the people involved have a right to our help is to obscure important moral differences. We stand in different

relations to these people and to their hunger, and it also seems reasonable to say, at least other things being equal, that our *priorities* should be determined accordingly.

No doubt it is true that, as we have come to establish a single world-wide economic system which affects everyone, more and more of our relations to distant others fall under the rubric of justice. And no doubt it is true that this began to happen long before the people of imperialistic nations acknowledged that it was so. Pogge is correct when he says that the language of rights seems appropriate when we understand the problem of hunger as the product of an unjust global economic system. But this does not support his case that moral philosophers can just as well use the language of rights as of obligations. It reflects a change in the world. And although ignoring the moral consequences of this change is one danger, overstating them is another. The fact of a single world economy does not imply that everyone exists at the same moral distance from everyone else. The world may now be one, but it has not collapsed into a point.

So far, really, all I have been doing is making a plea that moral language be used in a way that preserves and illuminates distinctions rather eradicating and obscuring them. This is a general reason for protesting Pogge's project of translating different moral idioms into one another. But there is also a more distinctly Kantian point at stake here. Starting with human rights and the world political situation and moving down, so to speak, to particular cases may conceal the fact that the moral world is constructed from where the individual stands, starting with her relations with her neighbors and working outwards. The possession of rights is not a metaphysical fact about people but rather something we learn to accord one another as we learn to interact morally. But of course for a Kantian the point from which the moral world is constructed is still deeper in. The moral world has its foothold in the individual's acknowledgement of the human plight: that we are autonomous but finite, free yet dependent on others who are free and dependent too. Your first thought, having grasped this fact, cannot be that it gives you rights as against others. It must be that it gives you obligations. I am autonomous and so I must give myself laws: that is where everything starts. In Kant's theory it is the

self-obligating agent who brings morality into the world. To this extent O'Neill's point about the priority of obligation can be given a foundational interpretation after all.

II. Universalization

The next part of my comments will be more brief. I want to discuss some of the problems Pogge uncovers in O'Neill's method of constructing duties. Pogge suggests that instead of asking O'Neill's question whether all could be guided by a given principle, we should ask whether all could *be free to* be guided by the principle. This formulation, he claims, solves a number of problems to which O'Neill's formulation gives rise, and which he illustrates by example. I want to start by observing that several of Pogge's examples either arise or seem to arise from two domains of human action which create problems for almost any version of the universalizability test, namely, the domains of coordination and competition. Coordination gives rise to problems for universalizability tests because certain resources and tasks must be shared out if human life is to be managed efficiently. People must divide up work and take turns with resources, and therefore it is sometimes essential that we should act differently from others. To borrow an example from Barbara Herman, from whom I have learned about these questions, you can make it your maxim to go to the tennis court at an unpopular time in order to have it to yourself, despite the fact that not everyone could act on this maxim without undercutting its aim. Competition gives rise to problems for universalizability tests, as well as for unanimity theories of goodness, because its apparent aim is to occupy a unique position - to be the winner, the best, or whatever - and this is not something that everyone can do.

The first case I will discuss is not obviously one of coordination, although I think that it turns out to be. Consider Pogge's question whether there is a duty to have children. If everyone followed the principle of remaining childless, the human species would die out. Pogge says that if we must will the continuation of the species, as the locus of morality, the principle of remaining childless will fail O'Neill's test. Yet it seems too strict to say that there is a duty to have children, things being as they

are. There are more than enough people who want children, so why should those who do not want children be saddled with this duty? However, if we ask whether everyone could be free to remain childless, we get the right answer, because enough people will have children even if they are free not to. I do not think that this analysis captures the important moral features of the case. Pogge arrives at his criticism of O'Neill by appealing to the necessity of willing the end of the continuation of the species. Clearly this is the morally important point: willing this end is what everyone must do. Then, recognizing that the continuation of the species is a species-wide cooperative project, what exactly one does will indeed depend on what others do and would like to do. In other words, the issue turns out to be one of coordination. What Pogge's analysis leaves out, initially, is the fact that, if the end he cites is indeed a duty, one is not free to remain childless if too many other people either cannot or will not have children. Pogge proposes to explain this fact by saying that if the preferences of the majority changed, so that not enough children were born, the duty to have children would then come into force. We would have a *new* duty, a duty to do something we previously were free not to do. To me it seems far more sensible to say that we have had the same duty all along, the duty of contributing to the continuation of the species. This is an imperfect duty, and exactly what or how much one contributes ordinarily depends on circumstances, including the majority's preferences. Like other imperfect duties, the duty to ensure the continuation of species can require action in emergency circumstances. Perhaps one must have children if no one else can do it, just as one must help the endangered if no one else is around to help. The other thing that this analysis captures which Pogge's leaves out is that even as things are those of us who remain childless must do our bit, whatever that might be, to ensure the continuation of humanity. The health and education of children is the concern of everyone, not just the responsibility of their parents.

Another of Pogge's examples concerns O'Neill's argument against slavery, which is that it cannot be universalized because if everyone were a slave there would be no one with property rights and so no one who was a slaveholder. According to Pogge, this fails because an exactly parallel

argument could be made against being an employee. If everyone were an employee no one would be an employer and so no one could be an employee. But in fact it's fine if some people want to be employees, because enough people want to be employers. Pogge's test is supposed to surmount this difficulty: people can be free to be employees. But this argument does not show that slavery is all right. To show this, Pogge switches to the argument against being a slaveholder: if everyone were free to be a slaveholder, there would not be any slaves, since the position of slaveholder is so much more desirable that everyone would take it. The switch in perspective makes his argument a little hard to follow. Can everyone be free to be an employer? Suppose enough people wanted to be boss that there was not enough employees to go around? Would this make being an employer wrong? However exactly it is supposed to work, the trouble with Pogge's argument in this case is that it makes it sound as if slavery is wrong because it presents a coordination problem that cannot be solved without the resort to force. But most employees are not exactly voluntarily employed. They must work to live, and in one sense their employers are taking advantage of that fact. The relevant difference here is that, at least if the economy meets certain conditions of justice, employers and employees *can* understand their relationships, however superficially contentious, as based on a form of cooperation beneficial to and shared by all. Whereas slaveholders and slaves cannot so understand their relationships; theirs is a relationship of victimizers and victims. In my view, a universalization test cannot successfully discriminate these cases unless it takes into account the *purposes* of being (say) a slaveholder or an employer. When you dominate others because that is your role in a system that benefits all, you are doing something fundamentally different, from a moral point of view, than when you dominate others in order to get the better of them.

Of course not every employer sees himself or herself as playing a role in a cooperative system that merely happens to take a competitive form. Many employers have as their purpose making more money than other people, and achieving power and domination over them, much as slaveholders do. Pogge thinks that the desire to overreach others is innocuous so long as the competitive activities in

which the overreaching takes place are voluntary. Like O'Neill, and while we are at it, Plato and Aristotle, I believe that the desire to overreach, *pleonexia*, is fundamentally wrong. Competition must be in the service of larger and more shareable goals: an exciting game, an efficient market economy, an improvement in excellence for all concerned. If Achilles' interest in being the best is not subsidiary to the more shareable desire to be as good as possible, then Achilles has as much reason to undercut his competitors as he has to improve himself. After all there is more than one way to be the best. You can try to make sure that no one else is any good. Achilles' father, when he advised him always to excel, did not have this method in mind.

The moral of these somewhat rambling remarks, if there is one, is this. Almost every version of the universalizability test runs into problems over the issues of coordination and competition. And part of the reason is that in these domains we perform necessarily exceptional actions which can be done in different spirits. You can see your actions as part of a larger and shared system, or as providing yourself with unique and unshareable advantages. It is possible that universalizability tests cannot discriminate finely enough without some attention to the attitudes and purposes that are involved when we act in ways in which not all could act. If this is true then neither O'Neill's version of the universalizability test nor Pogge's will succeed as they stand. Perhaps then it is fortunate that the supply of philosophers willing to work on the problem of universalizability never seems to run out. It means that, without violation of duty, at least some of you may act on the principle of being free to work on something else.