Rawls and Kant: On the Primacy of the Practical

Christine M. Korsgaard, Harvard

Our task today is to talk about Rawls's influence on Kant scholarship. Influence of this kind can be either methodological—an influence on the way we approach the subject—or substantive—an influence on the way we interpret the content of the view. Since I believe that Rawls has influenced Kant scholarship in both of these ways, I shall have something to say about each. But having said that, let me add a disclaimer: these two kinds of influence are less easy to distinguish than one might think, especially on the conception of the subject which I am about to describe.

Let me start with the methodological point. In Rawls's classes his students were exposed to an attitude towards the history of philosophy, and especially the great figures of the philosophical tradition, that I know I was not alone in finding exhilarating. I can best characterize this attitude by contrasting it with another. It is still common in some circles to make a rather strict separation between "doing philosophy" and "doing the history of philosophy." According to the view that these two activities are clearly distinct, which I will call "the strict separation view," when you study a great philosopher historically, your aim is to reconstruct the philosopher's position as exactly as possible, or, if it developed and changed over time, to chart these changes as accurately as you can. The enterprise is conceived as a scholarly one, and, in a certain way, the philosopher's view is conceived as having frozen or solidified on the day when he died, and, to that extent, as having died with him. By that I mean that a dead philosopher's views are treated as a *theory,* a certain body of doctrine organized in a logical way, with some of the doctrines derived from or dependent upon others, and tensions and inconsistencies noted now only as blemishes in the final product. It is possible instead to see a dead philosopher's views, as I think we do see a living one's, as something more like a project or even a method, a way of approaching the apparently intractable questions of the subject. Doctrines are of course involved in a project or a method, but a method is not merely a set of doctrines. What I have in mind here is not simply that the philosopher's doctrines might be *applied* to the problems, as when one tries to work out whether something is right or wrong by applying the categorical imperative. Later I will try to describe what I think is involved in treating a philosophical problem in a distinctively Kantian way, but I am sure you all recognize some ways of proceeding as meeting this description. Confronted with an apparent dilemma, for instance, we might ask: if we suppose Kant is correct in thinking that we view ourselves from two standpoints, might that shed light on why this dilemma exists? Or might the dilemma perhaps be recast as a sort of antinomy, with the existence of the two standpoints coming in as a solution? And where the phi-

losopher is, like Kant, a deeply systematic philosopher, treating his work as a method will also involve comparing similar problems which arise in different realms. For instance we might say: Kant thinks that there are deep problems about the unity of the thinker as such, which are resolved in such-and-such a way: here is a problem about the unity of the agent as such, so perhaps it is parallel, and perhaps a parallel solution can be found.

When we think of a philosopher's views in this way, they become something living, and, to put the point a bit bluntly, usable. Kant generated a system that was not only a set of doctrines but a method, and used it to work through a certain range of problems; you study the system by retracing his tracks until you get the hang of the way he did things, and then you can keep going on in something like the same way yourself, sometimes into territory he never got near. One difference this makes is worth noting. Proponents of the strict separation view tend to make a lot of the distinction between what is "Kant's view" and what is merely a "Kantian view." The rough idea is that "Kant's view" is what he said or what follows logically from what he said, while a merely "Kantian view" involves a revision of one or more of the fundamental premises of his theory. When you look at a philosopher's work in the way I've been trying to describe, this distinction tends to blur almost into nonexistence. To the extent that a philosopher's work is a method, it becomes clear that using it is something which can be done well or badly, not only by later disciples who take it up, but by the person who gave it his name. So this attitude towards the subject can make you rather fearless about saying things like "Kant really didn't say what he meant here" and these kinds of remarks become absolutely *continuous* with moments in which you undertake to pronounce on what a Kantian view is of a question which Kant himself could never have thought about at all.

Now the other side of what I am calling the strict separation view concerns the way in which philosophers who style themselves as "doing philosophy" rather than "doing history" approach and make use of history. Strict separationists who think of themselves as "doing philosophy" tend to treat historical figures merely as having given their names to positions which are thought of in the most nakedly schematic ways. The artificially sharp distinction between "Kant's views" and "Kantian views" comes in here too, and I think in a rather pernicious way, for it is used to excuse some rather violent historical distortions. There are still people around who will refer to a contemporary philosopher as "Kantian" simply on the grounds that he deploys a universalizability principle. The contemporary philosopher may think that the only valuable thing in the world is agreeable experience; and freedom, culture, and respect for humanity may play at most a derivative role in his thought; but no matter: he thinks we should ask "what if everybody did that?" and so he is a "Kantian." Strict separationists think of the history of philosophy as a kind of warehouse of such schematic positions, which can simply be removed from their original contexts

the way an object packed away in a box can be removed from its styrofoam packing pieces. You can just blow off the dust and put the thing in a new place.

Now I think that the view of a great philosopher's work as a method for solving philosophical problems does *not* license this kind of distortion. Especially when the philosopher is as systematic as Kant, you work in the awareness that a change or an addition in one place may redound throughout the system, and isn't to be undertaken lightly. And you also work in the awareness that the original author of the system probably said a great deal of what he said for reasons traceable to the systematic features of his work, and that you had better be aware of those reasons when you tinker with it. Sometimes I am tempted to think of a philosophical system as a kind of complex and intricately structured machine, like an engine. You can use it on problems that its originator didn't use it on, and you might occasionally revise it to make it work better, but if you are going to do this sort of thing well, you had better make yourself an absolute master of the original schematics, and have a good understanding of how the parts all work together, or you are likely just to break it. And this means that studying a great figure in the history of philosophy with a view to the solution of contemporary problems is not only consistent with, but absolutely requires, an attention to the *text* every bit as *reverent* as that called for by the most scholarly of enterprises.

Now I don't mean to say that John Rawls ever announced this as exactly his view of how one might proceed to learn from the history of philosophy, and I don't even know to what extent he would agree with the things I'm saying. But in his courses in the history of moral philosophy, Rawls always started by saying: "We are not going to criticize these thinkers, but rather to interpret their positions in ways that make the best of them, and to see what we can learn from them." And I think what he wanted us to see is the connection between the two sides of that formula—that it is *by* making the best of a philosopher's position that you can learn the most from it. The attitude shows in Rawls's own work, although it is more explicit in his earlier papers, which contain lengthy footnotes remarking on the ways in which Hume, Kant, Mill, and Sidgwick treated the issues with which he is concerned. Looking at these footnotes makes it perfectly clear that Rawls learned a lot of what he knows about these questions by thinking hard about what his predecessors knew. And what I take to be a flat undeniable fact—namely, that Rawls *gets somewhere* with the philosophical problems which *he* undertakes to solve—is bound to recommend his methods to the rest of us. And it has.

So much for the more general methodological point. I now want to take up my more substantive point. I want to suggest that both Kant and Rawls see philosophy as being, in a very deep way, a practical subject. I'm going to try to explain, and then illustrate, what I mean by this claim.

Philosophers who conceive the subject theoretically characteristically suppose that philosophical problems arise from gaps or unclarities in our theories—from something we don't know. Although it never happens exactly this way, such philosophers aspire to solve these problems by arguing from something like unassailable premises, and by showing that something else connected to these premises may be established—either as their logical consequence, or as their most natural interpretation, or as the best explanation for them. The philosopher's proposal is supposed to win the reader's assent by showing her first that the premises really *are* unassailable, and then that this commits her to the rest of the philosopher's view.

Philosophy conceived practically by contrast takes its start not from a premise but from a plight. Philosophical problems are seen to arise not from a gap in our theoretical views but from a problematic position in which we find ourselves. The philosopher's job is to describe that position and the problem to which it gives rise in a way that enables the reader to recognize it as *her own* problem, and he wins the reader's assent by exhibiting his proposal as the best or the only solution to it. The method is not unique to Kant and Rawls. Hobbes works this way in the *Leviathan,* and Descartes in the *Meditations.* But it is particularly central to the work of Kant and Rawls.

I say that it is particularly central to the work of Kant, because I think that this approach is actually called for by the correct interpretation of Kant's doctrine of the two standpoints. As I understand that doctrine, it goes like this: In one sense the world is given to us, it *appears* to us, and we are passive in the face of it. We must therefore think of the world as generating the appearances, as giving them to us. The world insofar as it appears to us is phenomenal; the world insofar as it *actively* generates the appearances is noumenal. We can only *know* the world as phenomenal, that is, insofar as it is given to sense, but we can *think* of it as noumenal. So there are not "two worlds," but rather one world which must be conceived in two different ways. And all of these points apply above all to ourselves. When we view *ourselves* as phenomena, we regard everything about ourselves, including inner appearances such as thoughts and choices, as parts of the natural world, and therefore as governed by causal laws. But insofar as we are rational, we also regard ourselves as *active* beings, who are the authors of our thoughts and choices, and therefore as noumena. We do not regard our thoughts and choices merely as things that *happen* in us; rather, thinking and choosing are things that we *do.* And from this standpoint, we can recognize laws that govern our mental powers in a different way than the laws of nature do: laws for the employment, for the use, of these powers; laws that show us how thinking and choosing must be *done.*[1]

There may be problems with this way of looking at the human situation, although I don't think it commits us to a belief in any mysterious form of supersensuous existence. It does, however, have important implications for the way

that we approach philosophical problems. It means that what we are seeking in philosophy is fundamental laws for the *employment* of our mental powers, laws to *guide* our proceedings, not laws whose primary aim is to explain some realm of phenomena. The basic task of moral philosophy, for Kant, is to answer the question "What should I do?" This task is set for us by our practically rational nature, which brings with it both the capacity for and the necessity of *choosing* our actions. Choice is our plight, our inescapable fate, as rational beings. The project of the critical moral philosophy is to determine what resources we can find in reason for solving the problem which reason itself has set for us.

With this as background, I want to compare two central arguments in the work of Kant and Rawls. Consider the opening argument of the third section of the *Groundwork.*[2] It is the announcement of a problem. Kant begins by defining a free will as a rational causality which is effective without being determined by any alien cause. Anything outside of the will counts as an alien cause, including the desires and inclinations of the person. The free will must be entirely self-determining. Yet, because the will is a cause, it must act according to some law or other: a lawless cause is a kind of contradiction. Alternatively, we may say that since the will is practical reason, it cannot be conceived as acting and choosing for no reason. Since reasons are derived from principles, the free will must have a principle. But because the will is free, no law or principle can be imposed on it from outside. Kant concludes that the will must be autonomous: that is, it must have its *own* law or principle. But now we seem to have a problem: for where is this law to come from? If it is imposed on the will from outside then the will is not free. So the will must make the law for itself. But until the will has a law or a principle, there is nothing from which it can derive a reason. So how can it have any reason for making one law rather than another? And indeed the problem is in a way even worse than that. For it looks as if the free will, by imposing some law or principle upon itself, must restrict its own freedom in some arbitrary way.

Now the problem which Rawls confronts is almost exactly like this. Political philosophers have long been aware that there is a kind of paradox at the very heart of liberalism. The problem emerges most starkly if you imagine someone trying to argue in favor of instituting a liberal regime in a nation whose culture and beliefs are not liberal. Imagine a nation in which there is a state religion which people must practice. This religion leads to illiberal policies: it assigns a subordinate role to women, say, or involves a caste system. It is, however, accepted by the population. Anyone who wanted to argue that more liberal political policies should be instituted in this society would face an intractable problem, for it is an essential tenet of liberalism that the government should be acceptable in the eyes of its people. If liberalism is the doctrine that you can't push people around in the name of what you think is right, then liberals themselves are committed to the view that *they* can't push people around in the name

of the doctrine that you can't push people around in the name of what you think is right. To put the point more simply, we cannot tyrannize over others in the name of liberalism and still be consistent liberals. By its own criterion, a liberal state must be acceptable in the eyes of its citizens. So no matter how passionately one was convinced that a liberal regime is the best one or the right one for human beings to live under, one could not consistently take that as a reason for forcing a liberal regime on an unwilling population.

Now Rawls is not, in this sense, trying to justify the liberal state. That is, he is not trying to give arguments that would show that there are grounds for forcing a liberal state on an unwilling population. But he is concerned about a parallel problem which arises when we try to justify policies *within* the liberal state, for even within the liberal state, we must use the coercive mechanisms of the state to enforce liberal policies. Since liberalism claims that political institutions and policies are justified when they are acceptable in the eyes of the citizens, we must be able to offer reasons in support of these coercive policies which are acceptable to all the citizens. To illustrate: suppose a society contains a majority and a minority religion, and suppose the majority wish to get their church accepted as the official state church. What reasons can they offer for this? They might say "Our religion should be the state religion, because our religion is the one true faith." This is not a reason that the minority can reasonably be expected to accept, since of course they don't think so. So in a liberal society, this will not be accepted as a good reason for coercive action.

Of course the majority are not just going to insist that they are right: they have certain arguments for their view—metaphysical, theological, and historical arguments—and they can marshal these and try to convince the minority that they are right. But the minority also have these kinds of arguments, and we may suppose that when all is said and done the disagreement remains. In a modern society, people hold different philosophical, theological, and metaphysical doctrines, which are reasonable in the sense that they need not involve the kind of obvious error or craziness that entitles us to regard people as mentally defective or emotionally deranged. These doctrines can be backed by arguments, but these arguments do not convince everybody or settle the question.[3] In the face of this fact it is inconsistent with liberalism to justify political policy on metaphysical or theological grounds, since such grounds may not be acceptable to all.

So now suppose that the majority say "Look, we want to use the coercive mechanisms of the state to ensure that everyone's soul is saved. Nothing is more important than that. Why shouldn't we be allowed to do this?" We cannot now turn around and justify liberal policies *themselves* on controversial metaphysical or philosophical grounds. For instance, we cannot consistently justify liberty of conscience by appeal to a controversial philosophical theory that says that there just are certain inalienable human rights, and liberty of conscience is one of

them; or by appeal to Kant's theory that autonomy is the supreme moral value; or by appeal to Mill's arguments for the utilitarian value of open discussion and experimental living. These may be excellent reasons for believing in liberty of conscience, if they are true; but they are controversial, and so they do not meet the criterion of being acceptable in the eyes of everyone. Our argument for liberty of conscience must *itself* be one that everyone can accept, if it is to justify the liberal policy in a liberal way. But this seems to leave us at a loss. How are we to make an argument which everyone accepts, in a society where there is apparently nothing that everyone accepts?[4]

These two problems, Kant's and Rawls's, have the same structure. What we are looking for are principles themselves, for we need reasons, ways of justifying our actions or our policies, and reasons are derived from principles. Yet the very structure of the situation seems to forbid us to choose any *particular* principles. The liberal's need to maintain neutrality among various competing conceptions of the good exactly matches the free will's need to maintain its own freedom. In each case, it looks as if the choice of any given principle must be arbitrary. In Rawls's construction of his problem, it looks as if the choice of any particular principle of justice must be based on an arbitrary preference for one conception of the good over others; in Kant's, it looks as if the choice of any principle of volition must involve an arbitrary restriction of the will's spontaneity. And the solutions proposed by Kant and Rawls take an exactly parallel form.

Kant's solution goes like this: The categorical imperative, as represented by the Formula of Universal Law, tells us to act only on a maxim which we could will to be a law. And *this*, according to Kant, is the law of a free will. To see why, we need only compare the *problem* faced by the free will with the *content* of the categorical imperative. The problem faced by the free will is this: the will must have a law, but because the will is free, it must be its own law. And nothing determines what that law must be. *All that it has to be is a law.* Now consider the content of the categorical imperative, as represented by the Formula of Universal Law. The categorical imperative merely tells us to choose a law. Its only constraint on our choice is that it have the form of a law. And nothing determines what that law must be. *All that it has to be is a law.* Kant concludes that the categorical imperative *just is* the law of a free will. It does not impose any external constraint on the free will's activities, but simply arises from the nature of the will. It describes what a free will must do in order to be what it is. It must choose a maxim it can regard as a law.[5]

Rawls's solution to his problem may be put in almost parallel terms. Rawls's two principles of justice tell us that all citizens must have equal basic liberties, and that our society must otherwise be designed so that it is to the benefit of everyone.[6] And *these*, Rawls might say, *just are* the principles of justice for a liberal society. To see why, we need only compare the *problem* faced by the

parties in the original position—those who are to choose the principles of justice—with the *content* of the two principles.[7] Echoing Rousseau, we might say that the problem faced in the original position is this: to find a conception of justice which enables every member of society to pursue her conception of the good as effectively as possible, while leaving each member as free as she was before.[8] The content of Rawls's two principles simply reflect this conception of the problem. So they describe what a liberal society must do in order to be what it is. We arrive at the principles of justice for a liberal society by thinking about the problem faced by a liberal society, just as we arrive at the categorical imperative by thinking about the problem faced by a free will. In each case, in fact, a sufficiently detailed and accurate description of the problem yields its solution. There are other points in common between the two arguments which I have just sketched. Both begin, although for rather different reasons, from the realization that we are in territory where theoretical metaphysics cannot help us. This is not just because we haven't got it, but because the problems we face, the problems we have to solve, are in any case practical rather than theoretical. Kant's view, as I understand it, is not just that we cannot have metaphysical knowledge about whether there are free wills, or what they would do if they existed. Even if we did, that wouldn't solve the problem posed by the fact that we must act under the idea of freedom, since what we must decide is *what to do*. And in Rawls's liberal state, we can't appeal to our metaphysical convictions about justice even if we have them, since the problem is posed by the legitimacy of disagreement about exactly such conceptions. Since the problems are practical, the solutions must be practical as well. And in constructing these solutions, both philosophers deploy practical conceptions of the person: as free to choose her own ends and reasons, in Kant's case, and as rational (that is, able to formulate, revise, and intelligently pursue a conception of the good) and reasonable (able to cooperate with others whose conceptions may be different) in Rawls's case.[9] This is not because these are the known metaphysical facts about persons, but rather because these conceptions of the person are in a sense foisted upon us by the problems we face. For Kant, the necessity of acting under the idea of freedom is a fundamental feature of the plight of reason; without that necessity, the problem of morality would not arise. For Rawls, in the same way, rationality and reasonableness are fundamental attributes of persons conceived as citizens of a liberal society, for unless people had both conceptions of the good they wished to pursue, and an ability to cooperate with those who have different conceptions, the problem posed by liberal society could not arise. In this way, both Kant and Rawls tackle problems conceived in a practical way by constructing a realm of practical concepts and arguments. In my view, both the reinterpretation of Kant's ideas in these terms, and, following Rawls, new employments of this method of solving philosophical problems, are projects worthy of further work.

Notes

1. Evidence for this interpretation of the two standpoints doctrine may be found in the Third Section of the *Groundwork of the Metaphysics of Morals*, especially pp. 451-452. There Kant argues that "even the commonest understanding" distinguishes appearances from things in themselves as a result of observing the difference between representations in the face of which we are passive and those in which we show our own activity. He then argues that we place ourselves among the things in themselves because of the pure spontaneity of reason, which in its production of ideas involves no element of passivity at all. The claim is clearly that we regard ourselves as things in themselves insofar as we are active. It is immediately after these remarks that Kant says that "a rational being...has two standpoints from which he can regard himself and know the laws of the *employment* of his powers...." (p. 452; my emphasis). The page numbers here are those of the Prussian Academy Edition (*Kants gesammelte Schriften*, Berlin: Walter de Gruyter & Company, 1902-.) found in the margins of most translations.

2. *Groundwork of the Metaphysics of Morals*, pp. 446-447.

3. In *Political Liberalism* (New York: Columbia University Press, 1993), Rawls calls this "the fact of reasonable pluralism." See p. 24n., 36ff.

4. This problem is much more clearly in focus in *Political Liberalism* than it was in Rawls's earlier book, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), and the modifications in the way Rawls presents his view in the later work are largely due to his increased appreciation of its depth and difficulty. Nevertheless, as I am about to argue and as Rawls himself believes, the strategy of the argument in *A Theory of Justice* does provide for its solution.

5. I defend this interpretation of Kant's solution at greater length in my essay "Morality as Freedom," in Yirmiyahu Yovel, ed., *Kant's Practical Philosophy Reconsidered* (Dordrecht: Kluwer Academic Publishers, 1989); which is also Chapter 6 in Korsgaard, *Creating the Kingdom of Ends* (New York: Cambridge University Press, forthcoming).

6. This is a deliberately general statement of Rawls's two principles, which he states with increasing specificity as he develops his view. For more exact formulations, see *A Theory of Justice*, pp. 302-303 and *Political Liberalism*, pp. 5 and 271.

7. For the idea of constructing a solution from the "original position" see *A Theory of Justice*, pp. 17-22, and *Political Liberalism*, pp. 22-28.

8. Rousseau says that the problem solved by the social contract is to "Find a form of association which defends and protects with all common forces the person and the goods of each associate, and by means of which each one, while uniting with all, nevertheless obeys only himself and remains as free as before." Quoted from p. 148 of *On The Social Contract* in Donald A. Cress, trans., *The Basic Political Writings of Jean-Jacques Rousseau* (Indianapolis: Hackett Publishing Company, 1987).

9. See *Political Liberalism*, pp. 48-54.