

## **Comment on Robert N. Johnson, “Kant’s ‘No-Accident’ Criterion of Moral Worth”**

**by Christine M. Korsgaard**

In the first section of the *Groundwork*, Kant claims that in order for an action to have moral worth it must be done from the motive of duty. The examples that follow contrast cases in which an agent does the right thing from natural inclination with cases in which the agent does the same thing from the motive of duty. Kant’s readers often wonder why he supposes that only actions done from the motive of duty have moral worth. One reason is suggested by Kant’s occasional claims that unless you act from the motive of duty, it is an “accident” that you act rightly.

Robert Johnson begins by offering a general account of the point of this “no accident” criterion. He says - I am quoting -

“what we want to know is whether it was the agent’s will alone, not happy circumstance, that was responsible for the rightness of her action. And we can know this only if we can be confident that her motive would have produced the right action had those circumstances been different.” (p. 3)

Johnson then contrasts two views of what it means to say it is “no accident” that you’ve acted rightly. The first view is that only the motive of duty produces right actions reliably. This view has been criticized on two grounds. What Johnson calls the “no-distinction objection” says the view is untrue. An intelligent, sensitive, sympathetic person who is never motivated by duty might produce right actions more reliably than someone who is capable of acting on the motive of duty but who is often weak. Or, to take a somewhat different case which I think Johnson neglects, this sympathetic person might do more right actions than a dutiful person who lacks intelligence or good judgment. What Johnson calls the “unreasonable demand objection” says that one may have done something morally worthy in a given case even if one’s motive, whatever it is, would not have been strong enough to produce the right action under stress of greater temptations. It is simply wrong to suppose that a morally worthy action can only be

done by a person of heroic virtue who would do the right thing under any temptation. Later Johnson mentions a third problem. A Kantian has no obvious reason to be interested in the question whether a motive reliably produces right actions anyway. Kant does not think that right actions ought to be maximized, or anything of that kind. The fact that a motive produces right actions reliably gives it only an extrinsic value, but the motive of duty, since it is expressive of a good will, is supposed to have unconditional or intrinsic value.

The second view of the no-accident criterion is Barbara Herman's. To say it is "no accident" that you did the right thing is to say that there was an internal connection between your motive and your action. You were interested in doing the right thing. This avoids the criticisms of the reliability view, but Johnson thinks there is a cost. As I said a moment ago, Johnson thinks the point of the no-accident criterion is to show that the agent's will rather than mere circumstance was responsible for the rightness of her action. And he thinks that this in turn requires that the claim that it was no accident that you did the right thing should imply the truth of certain counterfactuals, counterfactuals which state that you would have done the right thing had circumstances been different. But Herman's reading of the criterion does not seem to support any counterfactuals. An ephemeral interest in doing the right thing for its own sake might operate on a single unrepeatable occasion. The internal connection between motive and action would be in place, but the action would seem to be an accident all the same.

Johnson thinks this leads to a dilemma. Can we give an account of the no-accident criterion which supports counterfactuals without falling back into the reliability interpretation? His solution is to distinguish two kinds of counterfactuals: those which reveal a motive's reliability, and those which reveal its purity. Counterfactuals revealing reliability ask this question: Is this motive always strong enough to produce right actions? Counterfactuals revealing purity ask this question: if this motive did prevail in other circumstances, would it always lead the agent to do the right thing? Both ask whether, in other circumstances, the agent who had this motive would do the right thing, but the test for reliability reveals

something about the strength of the motive, while the test for purity reveals something about its content. The no-accident criterion, according to Johnson, attributes moral worth to those motives which would lead us to do the right thing in any circumstance, *if* they did prevail - that is, to motives which are pure. Since a prevailing sympathy can easily lead us to do the wrong thing, it fails this test. Kant's point, as Johnson understands him, is that *only* the motive of duty itself can pass this test, so only it has moral worth. As Johnson notes, however, Kant's argument so interpreted does not support the claim that only the motive of duty *as Kant understood it* - that is, the motive of acting from the categorical imperative - meets the test. Any form of principled or conscientious action will meet it. Kant's theory of the right is therefore left in need of independent justification.

A first point, which I hinted at earlier, is that Johnson seems to neglect the possibility that the agent's maxim might be pure and yet not produce right actions, not because the agent is weak and fails to act on the motive of duty, but because the agent is stupid and doesn't apply the categorical imperative correctly. This point may look picky, but I think it is not. We can see this by looking at the opposite case. Johnson mentions the possibility that sympathy may lead us to wrong actions because it leads us to focus on a person's condition rather than on her autonomy. One might argue, however, that a sufficiently intelligent sympathy would not make this error. Imagine a sympathetic person who believes that most people, whether they quite realize it or not, care deeply about their autonomy. He wants to give people what they really want, so his sympathy leads him to respect this concern. That kind of sympathy might produce right actions whenever it prevailed, and so meet Johnson's criterion. But it would not have moral worth. By contrast, wrong actions done under the influence of honest mistake that they are right do have moral worth. If this is so, the moral worth of a motive has nothing to do with whether it produces right actions, either reliably or purely.

My point is that Kant is not offering a no-accident criterion of moral worth. To see this, we need to keep in mind what the point of the first section of the *Groundwork* is. Kant opens with the claim that only a good will, and actions which are expressive of a good will, have

unconditional value. To say that these actions have unconditional value is to say that they have a value which is independent of what they “effect or accomplish” (394) - they are intrinsically good. Kant offers this to us as a distinctive feature of moral value: morally worthy actions have their value in themselves. Kant then goes on to assert - and I want to hammer a little on “assert” here - that the concept of duty includes that of the good will. (397) The reason I want to hammer on “assert” is that this makes it clear that Kant is not *arguing* that only actions from duty have moral worth. It is, in fact, one of his assumptions. It helps here to remember Kant’s statement, in the Preface, that he is “going to proceed analytically from ordinary knowledge to a determination of the supreme principle and then back again synthetically from an examination of this principle and its sources to ordinary knowledge ....” (392) In *Groundwork* I Kant analyzes ordinary rational knowledge of morality in order to discover its principle. The claims about the unconditional value of the good will and of actions from duty are his starting point. That is, Kant thinks that we all know that only actions done from the motive of duty have moral worth.

Now this by itself doesn’t settle the question whether Kant uses a no-accident criterion. Kant proceeds to try to find out what it is *about* actions from duty that gives them this special intrinsic value. We know that morally worthy actions are intrinsically good because of the way that they are willed - that is, because of the way they are chosen or decided upon. Kant wants to know what is distinctive about the way they are willed or chosen. Once we know how morally worthy actions are chosen, we will also know how we ought to choose our actions - that is, we will know what the supreme principle of morality is. That is the work of the argument of *Groundwork* I - to analyze the motive of duty in order to discover the supreme principle of morality. And you might think that this is where the no-accident criterion comes in - that what gives actions from duty their special intrinsic worth is that they are chosen in a way that makes their rightness non-accidental.

But this conclusion would of course be at odds with the text. For as Johnson himself says, actions which meet the no-accident criterion as he interprets it are not necessarily

actions done on the categorical imperative. And as I've just said, Kant's strategy is to find out what is distinctive about the way morally worthy actions are chosen, so that we will know how we should choose actions - that is, so that we will know what the supreme principle of morality is. Johnson thinks the analysis of good-willed motivation does not yield the categorical imperative. If Johnson is right, Kant's argument is a failure.

Let me try to sketch Kant's argument, briefly, to at least indicate why I think it is not a failure. In order to see what is distinctive about the way actions from duty are chosen, Kant contrasts them with similar actions which are done from direct inclination. After describing his three examples, Kant draws this lesson from them: "An action done from duty has its moral worth, not in the purpose that is to be attained by it, but in the maxim according to which the action is determined." (399)

The special value of a good-willed or dutiful action does not rest "in the purpose to be attained by it." This point follows from the fact that actions done from direct inclination and those done from the motive of duty may have the same purpose. The moral law, therefore, does not say: "Act for the sake of such-and-such purposes." Instead moral value rests in "the maxim according to which the action is determined." The moral law says "act on such-and-such a maxim." Okay, but what sort of maxim?

Here it helps to remember that when Kant says an action is determined according to a maxim, he is not talking about how it is caused. He is talking about what the agent takes into consideration when she chooses to do the action. And we already know, from the examples, that what the agent takes into consideration is not just the action's purpose, for there are different possible reasons for choosing a given action along with its purpose. You may, like the sympathetic person in Kant's example, choose to do an action that serves a certain purpose because you *like* it, because you have a natural inclination to pursue this purpose. Or you may, like the dutiful person, choose to do an action that serves a certain purpose because you think it is your duty to do so. Importantly, the trouble with the naturally sympathetic person - the reason his action lacks unconditional value - is not that he *wants* to help others only

because it gives him pleasure to do so. That's not so. The trouble with him is rather that he *chooses* to help others *only* because he *wants* to: when he decides to help others, he operates under what Kant calls the principle of self-love, that is, the principle of satisfying his desires.

By contrast, the dutiful person chooses to help others not just because she likes to but because she thinks it is a duty to do so - that is, to move ahead a little, because she acknowledges the *claims* that the humanity of others makes on her. Now to view an action as required of you, as making a claim on you, *is just* to view the maxim of doing that action as a law. Kant concludes that to be motivated by duty is to attend to the lawlike character -the lawlike form - of your maxims. This is the analysis of what is distinctive about good-willed or dutiful motivation that yields the categorical imperative, and, with it, Kant's theory of the right.

Why then all the talk about accidents? Here's why: If the maxim of an action has lawlike form, and the agent is moved by recognition of that fact, she views the action as being, in a sense, intrinsically valuable or intrinsically right - that is, she views it as something she has reason to do in virtue of the kind of action that it is. By contrast, someone who values, say, helping actions, only because he *likes* doing such actions views them as having only an extrinsic value, a value which they *inherit* from his desires. It *follows* from this that someone who does the right thing *only* because he likes doing it does not do it because it is right - and therefore, he does it by accident.