

A reply to Bencivenga, “Consequences in Kantian Ethics”

In Professor Bencivenga’s “Consequences in Kantian Ethics,” he offers a version of Kant’s ethics according to which the most rational approach to living one’s life is “to always imagine what might follow from one’s moves and to choose moves accordingly” (284), but according to which agents always nevertheless must be modest in their judgments about what they ought to do because the actual consequences of their actions might not turn out as they imagined. To get us to this conclusion, Bencivenga asks us to consider the character of Tuvia from the movie *Defiance*.

The details of Tuvia’s situation are not relevant for an assessment of Bencivenga’s interesting and novel argument. What is relevant is that after constructing two courses of action for Tuvia, Bencivenga argues that although we must assume that one option will emerge as the right one, the most that Tuvia (or anyone) can do “is hope that his move be the right one” (279). The idea here is that Tuvia cannot know whether the course of action he chooses is the right one (considered from the perspective of morality) until “the future is unpacked” (279). In other words, an agent cannot know whether an action is permissible according to Bencivenga until its consequences “play themselves out” (279). And because of this, our judgments about what we ought to do in any given situation might turn out to be wrong, for “elements of the situation which we regarded as unimportant may reveal themselves later to be decisive for whether our move was in fact made in the interest of reason” (284).

However, we are not entirely without recourse. That is, Bencivenga does think that agents have a way (one way) to determine whether a course of action was the right one: they must observe their own dispositions through time. If an agent “can recognize an improvement in his disposition,” then he justifiably may hope that his action was the right one (279). In other words,

although an agent cannot know with certainty whether his judgment about the permissibility of his action was correct until the consequences play themselves out in time fully, by observing his disposition over time he can at least get a better idea: if his disposition improves, then he has reason to believe that he behaved permissibly (this idea also seems to be underlying Bencivenga's discussion at 283: "the pattern of our actual behavior, which constantly unfolds during a war, will often not justify any hope in our own moral improvement").

As pointed out above, Bencivenga is putting this theory forward as a version of Kant's ethics. The claim discussed in the previous paragraph (that an agent has a justified hope that he has behaved well if he observes an improvement in his disposition) is made on the basis of the following passage from Kant's *Religion within the Boundaries of mere Reason*:

[A] human being who, from the time of his adoption of the principles of the good and throughout a sufficiently long life henceforth, has perceived the efficacy of these principles on what he does, i.e. on the conduct of his life as it steadily improves, and from that has cause to infer, but only by way of conjecture, a fundamental improvement in his disposition. (6:68)

However, inspection of this passage reveals that it does not say what Bencivenga takes it to say.

In this passage, Kant is telling us that one may infer an improvement in one's disposition from observing the efficacy of one's resolution to act on good principles. And presumably one observes the efficacy of one's resolution to act on good principles by observing one's behavior in individual instances. That is, the idea is that if an agent has resolved to act on good principles and then goes on to observe himself actually doing so, then he may infer an improvement in his disposition. In other words, knowledge of the permissibility of his behavior in individual instances is taken to be logically basic, and from this knowledge an agent may garner knowledge about his disposition. The idea is not (*pace* Bencivenga) that an agent justifiably may hope that he has behaved permissibly if he observes (logically basic) an improvement in his disposition.

I bring up this minor point about the passage from *Religion within the Boundaries of mere Reason* because underlying it is a much larger disagreement about uncertainty and the good will in Kant's ethics. At the beginning of part I of the *Groundwork for a Metaphysics of Morals*, Kant famously claimed that the goodness of a good still "would shine by itself, as something that has its full worth in itself" even if it were coupled with bad consequences (4:394). In other words, what matters for the purposes of ethics is, according to Kant, acting on good principles (cf. 5:63). Kant takes the rightness of rules to be logically basic. The goodness of consequences is derived from the rightness of rules. That is, to determine whether a given state of affairs is good, Kant thinks we must determine whether it was brought about by agents acting in accordance with right rules. Determining whether an action is good is not a matter of inspecting the consequences of that action to see whether they maximally realize some end to be produced (like happiness). Rather, to determine whether an action is good, for Kant, we must determine what principle the agent in question is acting on.

In saying this, I am disagreeing with Bencivenga's claim (discussed above) that to determine whether an agent's action is good we must wait for its consequences to unfold in time. As Bencivenga sees things, although agents might know what their maxims are, they do not know whether those maxims are permissible. But this puts him at loggerheads with Kant: Kant was somewhat pessimistic about our ability to discern our own (and others') maxims and inner motives, and he was quite optimistic about our ability to discern whether a given maxim is permissible. For example, in the *Critique of Practical Reason*, he claims, "the most common understanding can distinguish without instruction what form in a maxim makes it fit for a giving of universal law and what does not" (5:27; cf. 4:404 and 4:424).

Of course, someone might disagree with this. Someone might think that Kant was wrong to be so pessimistic about our ability to determine what our maxims are. And someone also might think that Kant was wrong to be so optimistic about our ability to determine whether maxims are permissible. But the point of calling on the passage above about the good will is to show where this disagreement would get us: not in a place where we would be appraising actions by appeal to their consequences or where we would be given hope that our actions were good only by noting an improvement in our own dispositions.

Disagreeing with Kant on either of these fronts would not get us there because we still would be left with a central tenet of Kant's ethics: that the goodness of a state of affairs can be determined only by appeal to the rule being followed by the agent who brought that state of affairs about. And whether that rule is permissible is determined by appeal to the Categorical Imperative, which does not ask about the consequences an action reasonably can be expected to have or actually will have in our world but rather about whether a maxim can be willed in a world in which it is a universal law (whether the maxim can be willed at the same time as a universal law).

I do think that consequences are important in Kant's ethics. For one thing, they inform what we are trying to decide to do: we cannot make sense of the classic maxim to tell a lying promise to get some ready money if the getting of the ready money is not conceived of as a consequence of telling the lying promise. The Hypothetical Imperative is important here. For another, they play a role in determining the circumstances that, as we are aware of them, inform what maxims are permissible in the continuing context of action: if I borrow your phone and accidentally break it, that does not obviously tell against me — but once that enters my

deliberative field, presumably I must adopt an appropriate maxim in my behavior to you in the continuing context of action, and presumably that will include making some sort of restitution. There is a complicated story to be told there. But the point is that if we are to tell it in Kantian style, then I think we must retain the focus on principles rather than consequences, on the meanings of maxim-based actions rather than on their results.

Samuel Kahn, Indiana University-Purdue University, Indianapolis