

Mill, Kant, Griffiths, and Williams on Raskolnikov's Crime

Christine M. Korsgaard

The discussants begin by addressing the question what Kant would say about the case. Kant's first formulation of the categorical imperative tells us to act only on a maxim—a principle of action—which we could will to be a universal law, that is, a principle upon which everyone acts. Kant means one that we could will to be a universal law rationally, without contradiction. There is controversy about how exactly we are supposed to contradict ourselves when we attempt to universalize bad principles.¹ Griffiths adopts the practical contradiction interpretation, according to which the contradiction emerges this way: A principle tells us to do a certain act in order to achieve a certain end. But if everyone attempted to do that act in order to achieve that end, the act would no longer work to produce the end—the principle would cease to be instrumentally rational. This interpretation works well in the case of actions which are defined in terms of conventions or institutions, or at least are dependent on the existence of conventions or institutions for their effectiveness. Kant argues, for example, that in a world in which everyone made false promises in order to achieve their ends, making false promises would cease to be effective as a way of achieving ends, because people who were constantly burned by having accepted false promises would cease to accept promises at all. The convention by which you can convince another to rely on you by saying the words “I promise” would be undermined, so that you could achieve nothing by saying those words. It is harder to make the interpretation work for actions which depend only on the laws of nature for their effectiveness. No matter how many people murdered wicked old women in order to stop them from doing harm, we might think, murder would still be an effective way to stop them from doing harm. But Griffiths argues that the broader purpose of the murderer's principle is to make people happier, and that in a world where people freely committed

¹ For a review of the various views of how the contradiction emerges, and longer versions of some of the arguments sketched in this paragraph, see Christine M. Korsgaard, “Kant's Formula of Universal Law,” in Korsgaard, *Creating the Kingdom of Ends*. Cambridge University Press, 1996.

murder in order to eliminate those whom they considered harmful, people would be insecure and therefore not happier after all. The maxim of killing harmful people in order to make everyone happier would, if universalized, no longer work, because everyone would be afraid of being deemed harmful.

An interesting feature of Griffith's argument is that it is right next door to one of the standard arguments that utilitarians use to defend themselves against the charge that they would have to endorse actions like killing the old woman as morally right. John Stuart Mill argued that people have rights that we are committed to upholding, even when we might do a lot of good by violating them, because rights protect a special kind of value. Mill says:

The interest involved is that of security, to everyone's feelings the most vital of all interests. All other earthly benefits are needed by one person, not needed by another... but security no human being can possibly do without; on it we depend for all our immunity from evil and for the whole value of all and every good, beyond the passing moment, since nothing but the gratification of the instant could be of any worth to us if we could be deprived of everything the next instant by whoever was momentarily stronger than ourselves.²

The value of almost any good thing we might have depends on our being secure in the possession of that thing, so security is a condition of almost all other value. Among the things we need security against is being killed by whoever is stronger than ourselves and deems us unworthy of life. It is especially interesting in this context to note that, although he is certainly not a utilitarian, Bernard Williams agrees with Mill about this. In a discussion of obligation, Williams says that obligation is grounded in:

² John Stuart Mill, *Utilitarianism* (Indianapolis: Hackett Publishing Company, 1979), p. 53.

... the basic issue of what people should be able to rely on. People must rely as far as possible on not being killed or used as a resource, and on having some space and objects and relations with other people they can count as their own... People need help but (unless they are very young, very old, or severely handicapped) not all the time. But all the time they need not to be killed, assaulted, or arbitrarily interfered with.³

So if Griffith and Mill are right, then utilitarians, Kantians, and Williams can all agree that allowing people to kill others whenever they deem it useful to do so would lead to effects that are harmful to our security. If an agent's aim in killing is to make people happier, and happiness requires security, this result would also be in contradiction with the agent's own intention in committing the murder.

But whether or not Kant could make the sort of argument that Griffith envisions, he has a better argument. In the *Metaphysical Principles of Right* (the first half of *The Metaphysics of Morals*), Kant argues that each of us has an innate right to freedom⁴ which includes a right against physical coercion,⁵ and therefore a right not to be killed. You cannot be free if you cannot determine what other people are allowed to do with your own body. And Kant makes it clear, when he discusses the second formulation of the categorical imperative, that we are morally obligated to respect people's rights.⁶ This is a better argument, because this argument brings out why killing the old woman wrongs the old woman in a particular way. Griffith's argument

³ Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge, MA: Harvard University Press, 1985), pp. 185-186

⁴ Immanuel Kant, *The Metaphysics of Morals*, translated by Mary Gregor (Cambridge: Cambridge University Press, 1996), 6:237, p. 30.

⁵ See for instance *The Metaphysics of Morals*, 6:248, p. 38.

⁶ Immanuel Kant, *Groundwork of the Metaphysics of Morals* (Cambridge: Cambridge University Press, 1997), 4:430, p. 38.

tells us why all of us, collectively, are wronged when someone acts on a principle that would bring about general insecurity if it were universalized. But it does not tell us why the old woman is wronged in a different way than the rest of us are, or why in general the victim of a crime or a wrongdoing is violated in a particular way. In fact, it seems plausible to say that what is most wrong about murder is the particular sort of wrong done to the victim.

This is a general advantage to Kant's second formulation of the categorical imperative, which tells us that we must never treat a human being as a mere means to an end, but instead must always treat her as an end in herself. Griffith says that this formulation "emphasizes the unconditional, absolute value of the individual." It seems plain that when we treat someone as a mere means to our ends who ought to be treated as an end in herself that we have wronged her, and not just done something that would be in general damaging to society if everyone did it.

Still, we might wonder whether the case fits the description of "treating the old woman as a mere means to an end." What we ordinarily mean by treating something as a means to an end is that we use it in a way that furthers our own ends—that we use it as a tool. I suppose we could make a case that robbing the old woman of her money is treating her as a tool. The student imagines doing a great deal of good to the poor with her money, which she plans to give to a monastery in order to secure prayers for herself. So if he did commit the robbery, the student would in effect have used the old woman to collect the money which is to be used for his own purposes rather than hers. (It is interesting to note here that the student envisions doing good to the poor in general with the money, rather than returning it to the customers from whom she collected it, even though he suggests that her niggardly practices as a pawnbroker have wronged them.)

But how is killing the old woman in order to stop her from doing harm treating her as a tool? Murder here is not exactly a way of producing good results; it is just a way of preventing

harmful ones. Perhaps a clearer case would help to show the problem I have in mind. Suppose that as I drive my car in a hurry somewhere, someone gets in my way, and I just run over him. It is not that running over him in any way furthers my end: it simply fails to allow him to present an obstacle to my endeavor to go as fast as I can. The claim that I should not do this cannot be derived from the idea that if I did, I would be treating him as a mere means to my end. Instead, we must first understand why it follows from his status as an end in itself that he should have presented an obstacle to my action. In a similar way, preventing the old woman from doing harm to her customers and her sister is not a really a means to securing their good, or a way of doing good; it is simply the removal of a source of harm. Before we know why we should not remove harm that way, we must know why it follows from her status as an end in itself that we should not just eliminate her. What these cases show is that “treating someone as an end in itself” must have a content that goes beyond “not treating her as a mere means.”

Intuitively, it seems plain enough that if you kill someone you fail to recognize her “unconditional, absolute value,” as Griffith puts it. But why is that exactly? It is not that we think that if something has value as an end, then then it should be preserved, and that if that value is unconditional and absolute, it should be preserved at any cost. We do not treat every valuable thing as a precious object. We respond to the value of cultural treasures by trying to preserve them, but we respond to the value of beautiful fireworks and blooming flowers by enjoying them while they last. In general, we respond to the value of different things in different ways. We respond to the value of pleasurable experiences by enjoying them ourselves and trying to bring them about for others; we respond to the value of beautiful things by appreciating them; we respond to the value of small children and pets by taking care of them. Why do we respond to old woman's value as an end in herself by refraining from killing her, even when her death might have value for other people?

The answer lies in the way we value people. According to Kant, what we do when we value people is respect their choices and promote their interests. One of the ways in which Kant's

theory is most obviously different from utilitarianism is the two views involve different accounts of what it means for people to have value. The utilitarian thinks that people have value—that is, that we should care about what happens to them—only because the things that happen to them have a value of their own. That is, the utilitarian thinks that happiness and misery, pleasure and pain, are inherently valuable and disvaluable states, and that people matter because they are subject to being in those states.⁷ But since it is the states themselves that really matter, the utilitarian measures the value of an event or condition by considering its effects on each person and aggregating those effects. It does not really matter which person gets the happiness or misery; what matters is just the total.⁸ This is why Williams says that the utilitarian is willing to balance anything against anything. If we can increase the total happiness by taking a source of happiness away from me and giving it to you, because you would enjoy it more, the utilitarian must be in favor of that, unless perhaps we can invoke Mill's security argument, or something like it, against doing that sort of thing.

Kant's injunction to treat each person as an end in herself forbids us to aggregate values in that way. To see why, it helps to think separately about the ideas of respecting people's choices and promoting their interests.

First consider respecting people's choices. This conception of what it means to value people as ends in themselves is reflected in Kant's view, mentioned above, that people have an innate right to freedom. When we value a person as an end in herself, we are committed to the idea that she should be free to live her own life, to choose her own actions and as far as possible to determine the purposes which her own life and actions will serve. As Griffith says, no one

⁷ The utilitarian also thinks that non-human animals matter this way. For an argument that animals also matter on the Kantian conception, as ends in themselves, see my *Fellow Creatures: Our Obligations to the Other Animals* (Oxford: Oxford University Press, 2018).

⁸ Or, in some versions of utilitarianism, the average.

else has a right to put the old woman's life and actions to work for their purposes instead of her own.

Now think about why the Kantian thinks we should promote people's interests. To derive the value of promoting people's interests *from* their value as ends in themselves is to say that it is people who have value in the first instance, and that their states only have value because they have value for people. According to the Kantian, your happiness, unhappiness, pleasure, and pain matter only because you matter. They matter because they matter to you. This is a reversal of the utilitarian idea that you matter only because you are a kind of locus for these states.

When we put these two thoughts together, we get the conclusion that as ends in themselves, people stand in a special relationship to the value of their own lives. The value of people's ends derives from the value of people themselves, but people have a right to determine what those ends will be. This means that the value of a person's life is above all its value for the person herself. To respect people as ends in themselves is to respect that fact about them. Because the old woman is an end in herself, the value that someone else might get from her life, say by using her as a slave, has no weight against her own choices about how to live her life. In just the same way, the value that someone else might get from disposing of her life has no weight against the value that her life has for her.

Notice that this gives us another way of hearing the student's remark that the old woman "doesn't know what she is living for." Perhaps the student is not merely denying that the old woman's life has value; perhaps he is trying to claim that her life has no value because it has no value for her. Of course, whether her life has value for her is not up to him to decide. If we hear the student's remark this way, we may think Dostoevsky is suggesting that the student is uncomfortably aware of what is wrong with his own argument: that it is inconsistent with the old woman's value as an end in herself.

© **Christine M. Korsgaard, June 2021**