

Response to Slote

Suppose no one could do otherwise. Or suppose at least that we came to believe that no one could do otherwise. What would the consequences be for ethics? Professor Slote suggests that it is possible to think of approaches to ethics that could survive in the absence of free will—or in the absence of a belief in it.

One such approach may be described as follows. Suppose that we were to adopt a "typical" utilitarian view of the assignment of blame. Slote suggests that a typical utilitarian view of the assignment of blame does not depend on the assumption that we have free will. But I am not quite clear as to what this view is.¹ Suppose we assume that the basic form of a statement for the assignment of blame is 'X is Y's fault.'² (Here, X is a state of affairs and Y is, of course, an agent.) What, according to the utilitarian theory of the assignment of blame, is the content of statements of this form? One possible reading would seem to be the following:

X is an unfortunate state of affairs and, for some group of people, it would maximize the general welfare if they were to do something unpleasant to Y and to describe their motive to Y and to the public as follows: We are doing this unpleasant thing to Y *because* Y brought about X.

Here is an example of this theory of the assignment of blame at work. Suppose that a certain town is in a panic because of a series of brutal murders. Lynch law is about to break out, and the police haven't a clue as to the identity of the murderer. In collusion with the courts, they arrest, frame, and speedily execute an unpopular local figure, thus calming the populace. According to the reading of the utilitarian theory of the assignment of blame we are now considering, the unfortunate scapegoat is to blame for the murder; it is his fault.

Copyright 1990 by *Social Theory and Practice*, Vol. 16, No. 3 (Fall 1990)

A second possible reading of the utilitarian theory of the assignment of blame is as follows:

X is an unfortunate state of affairs and *Y* brought about *X* and, for some group of people, it would maximize the general welfare if they were to do something unpleasant to *Y* and to describe their motive to *Y* and to the public as follows: We are doing this unpleasant thing to *Y because Y* brought about *X*.

The second reading differs from the first in just one respect: the second reading implies that *X* can be *Y*'s fault only if *Y* actually brought about *X*. Now this is no doubt overly restrictive, since we may sometimes want to hold people responsible for states of affairs they did not bring about—states of affairs they failed to prevent, for example. But let us ignore this problem. The important difference between the two readings is that the second implies that a certain causal relation must hold in fact (and not simply in the mind of the public) between the agent and the state of affairs alleged to be the agent's fault; in all probability, the required causal relation is more inclusive than *brings about*.

Here is an example of the second reading at work. A person of diminished mental capacity—a moron, say—seriously injures someone, owing to a combination of ignorance of his own strength and malicious goading from cruel bystanders. The police and the courts rightly judge that the moron will be lynched if he is acquitted on grounds of incapacity, so they conspire to deprive him of his legal rights (a court-appointed attorney agrees not to raise the point of his client's mental capacity at the trial), with the consequence that he is sent to prison for assault. According to the second reading of the utilitarian theory of the assignment of blame, the injury the moron was goaded into causing was his fault.

On either reading, the utilitarian theory of the assignment of blame seems plainly wrong. (Perhaps the second reading is not quite so plainly wrong as the first. It is, however, not evident to me on what ground a utilitarian could consistently prefer the second reading to the first.) Slote is well aware that many people will share my reaction to the implications of the utilitarian theory. He argues, however, that inconsistencies lurk in our everyday moral conceptions, and that we should therefore be suspicious of the moral intuitions that tell us that the utilitarian theory is "plainly

wrong." Moreover (Slote contends) the utilitarian theory neatly yields consistent results when it is applied to the cases that lead us to make inconsistent moral judgments when we rely on our untutored moral intuitions.

What are these inconsistencies? They are revealed, Slote says, when we examine our reactions to cases involving "moral luck." In a nutshell, the charge against our everyday moral intuitions is this: they lead us to say both that an attempted murder—one thwarted by chance—is not as bad as successful murder, and also that an attempted murder is just as bad as a successful murder.³

I come to these arguments as an innocent, for I am almost entirely unfamiliar with the rather large body of literature on the topic of moral luck. But perhaps the perceptions of an innocent will be of some interest. My naive and uninformed reaction to Slote's arguments is that I am not at all sure that reflection on moral luck reveals any incoherence in our everyday moral thinking. Let us look at some cases.

Case 1. Alice proposes to murder Beatrice, whose heir she is. When Beatrice is out of the room, Alice slips poison into Beatrice's tea, and Beatrice dies.

Case 2. Clara proposes to murder Delia, whose heir she is. But Delia is suspicious, and substitutes salt water for the poison. When Delia is out of the room, Clara pours the salt water (believing it to be the poison) into Delia's tea. When Delia returns and tastes her tea, her suspicions are confirmed.

Now what are the inconsistent reactions to these cases that are supposedly dictated by our ordinary moral intuitions?

I'm sure that all of us think that Alice and Clara are equally bad persons, at least insofar as their moral characters are revealed by the two stories.

I'm sure all of us think that the consequences of Alice's act (insofar as we can judge them from the story) are worse than the consequences of Clara's act.

I'm sure that all of us think that what Alice did is wrong and that what Clara did is wrong.

Perhaps the inconsistency is this: we have a tendency to think that the degree of wrongness of the two acts is the same, and we

have a simultaneous tendency to think that what Alice did is "wronger" than what Clara did.

This suggestion is promising, but there is some question as to whether it is intelligible. Does the adjective "wrong" in fact have a comparative degree? Isn't "wrong" in this respect like "against the law"? Isn't it the case that, just as one act can't be more against the law than another, so one act can't be "wronger" than another? Isn't a wrong act an act that is in violation of the moral law, just as a crime is an act in violation of the criminal law? While there may be something to this objection, there would seem to be a plausible response to it: "Although one act cannot be more against the law than another, one act can be a more serious offense against the law than another. Thus, murder is a more serious offense than petty theft, although both act are equally against the law. And in this case, positive law and the moral law render parallel judgments. Morality tells us that murder and theft are both, without qualification, wrong; but morality also tells us that murder is a more serious wrong than theft, a more serious or deeper offense against the moral law than theft. It is this concept, the concept of a more serious wrong, that is needed to describe the inconsistencies embedded in our ordinary moral convictions. When they are subjected to a certain line of Socratic interrogation, our ordinary moral convictions tell us that an attempted murder and an otherwise identical successful murder are equally serious moral wrongs; but they also tell us (in response to a different line of interrogation) that an attempted murder is a less serious moral wrong than an otherwise identical successful murder."

The first "line of Socratic interrogation" consists in leading the subject to assent to a kind of argument of which the following is a typical example. Consider Alice and Clara. What went on in their minds was—we may so stipulate—identical. The fact that their states of mind had different consequences in the external world is a matter of sheer luck, and luck is morally irrelevant. It is only what is done in the mind (or in the heart, as people used to say) that is really subject to moral judgment, for it is only what is done in the mind that is fully understood by the agent and fully under the agent's control. Indeed, an external object can be understood and controlled only insofar as it can be represented in the mind.

The "philosophical" language in which this argument has been cast should not be taken to imply that the argument represents a point of view which is the invention of philosophers. This point of view was, for example, an important part of the moral teaching presented in the Sermon on the Mount: "You have heard that it was said to the men of old, Thou shalt not commit adultery. But I say to you, whosoever looketh upon a woman to lust after her hath already committed adultery with her in his heart."

I think that it is pretty obviously true that most people can be brought to assent to the thesis that an attempted murder and an otherwise identical successful murder are equally serious moral wrongs. But what is the evidence that they can also be brought to assent to the thesis that an attempted murder is a less serious moral wrong than an otherwise identical successful murder?

There is the fact that the law distinguishes between murder and attempted murder, and regards the former as a much more serious offense. But it is not clear what we should conclude from this, for it could have all sorts of explanations that have nothing to do with our moral convictions. I'd like to hear from lawyers, judges, legislators, and police officers on this matter. What would the legal and social consequences be if, for every crime, the penalty for attempting that crime were the same as the penalty for committing it?

In point of fact, I am not sure what the evidence for this thesis is. But let us suppose that we do have some tendency to think that an attempted murder is a less serious wrong than is an otherwise identical successful murder. And let us suppose that this tendency is not confined to judgments about murder, but rather applies to any sort of moral wrong whatever. And let us grant that we also have a tendency to judge that an unsuccessful attempt to commit a moral wrong is just as serious an offense against morality as an otherwise identical attempt that, owing simply to chance, succeeds. It follows that our moral beliefs are inconsistent. (Or, at any rate, it follows that, for some moral judgments, we have a tendency to make those judgments and also have a tendency to make their denials. This may not be quite the same thing as having inconsistent moral beliefs, but I shall not pursue this point.) It is, however, an old philosophical discovery that an inconsistency in

a set of beliefs may be relatively superficial. That is, a set of beliefs may be inconsistent and yet be such that it can be turned into a consistent set by some adjustment that, intuitively, we judge to be of little importance. Other sets of beliefs may be radically inconsistent. Any revision of a radically inconsistent set of beliefs that yields a consistent set of beliefs is analogous to major surgery: our intuition is that the new, consistent set of beliefs must represent an important change in the believer's opinions. Thus, if Jane believes that God created everything, and also believes that God is uncreated, her beliefs are inconsistent. But they can easily be rendered consistent: Jane need only replace the former belief with a belief that God created everything besides Himself. If, on the other hand, John believes that inequality between the sexes is a consequence of social conditioning, that inequality between the sexes is a feature of every culture, that anything that is a feature of every culture is biologically determined, and that nothing is both biologically determined and a result of social conditioning, his beliefs are probably radically inconsistent.

Superficial and radical inconsistency are the extremes on a continuum. In between are inconsistencies of various degrees of seriousness. It seems to me that we should not be justified in abandoning our ordinary moral convictions and adopting something so deeply at variance with them as the utilitarian theory of the assignment of blame unless our ordinary moral convictions were at least seriously inconsistent. Something like the Principle of Minimum Mutilation seems in order here. Let us therefore examine the possibility that the inconsistencies in our ordinary moral convictions (always supposing them to exist) can be removed by a revision that is minor compared with the revision that would be entailed by adopting the utilitarian theory of the assignment of blame.

If we are inclined to judge that an attempted murder is a less serious offense than an otherwise identical successful murder, this may be because we have a natural tendency to desire revenge. And it may be that human psychology is such that we find it easier to "forgive and forget" after a mere attempt ("Well—no harm done") than after a success. It is, of course, true that to attempt to murder, say, a child, is to commit an offense against that child (and against

the child's parents), and no doubt many people would attempt to exact vengeance from someone who attempted to murder them (or their children) but who succeeded in doing no one any harm. And yet it does seem plausible to suppose that a mother who acted to avenge the attempted murder of her child, would, if she could, avenge the successful murder of her child by some act more drastic than the one she deemed appropriate in the case of the attempted murder. Perhaps it is because our desire for revenge is weaker in the case of an attempted murder than in the case of an otherwise identical successful murder that we are inclined (if we indeed are inclined) to judge the former to be a less serious offense than the latter. And perhaps there are other natural human feelings that contribute to this inclination. It may well be that I should be fairly comfortable sitting beside Clara, who attempted to murder someone and was thwarted by the merest chance, and that I should be quite uncomfortable sitting beside her if she had succeeded in her attempt.

If this is so, then let us revise our ordinary moral beliefs by discounting those of our inclinations to make moral judgments that are based on a desire for revenge or feelings of revulsion against transgressors, and not by simply scrapping the whole lot (and that is what adopting the utilitarian theory of the assignment of blame would come to).

To sum up what I have said: Firstly, it is clear that the utilitarian theory of the assignment of blame runs deeply contrary to our ordinary moral convictions, and, secondly, if considerations of "moral luck" do indeed show that there is any inconsistency in our ordinary moral convictions, they certainly do not show that this inconsistency is so radical as to require the desperate remedy of adopting the utilitarian theory of the assignment of blame. (I have suggested a more modest remedy. If that more modest remedy is not right, we should look long and hard for other modest remedies before turning to the utilitarian theory.)

Slote also suggests a second approach to ethics (besides adopting the utilitarian theory of assignment of blame) that could survive in the absence of a belief in free will: we might consider adopting a "virtue ethic." As I understand the notion, a virtue ethic is an ethical theory in which the only ethical judgments allowed

are those that pertain to a person's character (or long-term dispositions to behavior)—or, at least, a theory according to which all ethical judgments must ultimately be reducible to judgments about character.

In this matter I can only record my conviction that a pure virtue ethic must leave something out.⁴ There are things that I have done that I deeply regret. (I will not tell you about any of them, for the simple reason that I am so ashamed of them that I do not want anyone to know what they were.) Some at least of these things were not an expression of my character. With respect to acts of those sorts, my character was yet unformed. If I had continued to perform acts of those sorts, a tendency—perhaps eventually an irresistible tendency—to perform them would have become a part of my character: my character would have become, in those respects, vicious. Unfortunately, there have been other acts that I regret, acts that, to my cost, I repeated (by my own free choice) till a tendency to perform them became a settled part of my character. I have, therefore, freely chosen to have a character that is, in certain respects, vicious, and I am responsible for these vicious aspects of my character.

I do not mean to imply that one is responsible for every aspect of one's character. No doubt one's character is to a significant degree determined by factors outside one's control, factors like one's genetic makeup and the characters and social resources of one's parents. I am nevertheless convinced that I could have a significantly better character than I do, and that I am responsible for the fact that it is not significantly better. And I am convinced that most physiologically normal people who have not had "pathological" upbringings could have had better characters than they do.

Here is what, in my view, a pure virtue ethic must leave out: There is a standard of behavior that exists antecedently to and independently of human character, and any reasonably normal human being—at any rate, any human being who is in a position to read these words—could have so acted that his or her character conformed better to this standard than it in fact does; every reasonably normal human being is therefore responsible for some of the defects in his or her character.

This, of course, is only my opinion, and doubtless many philosophers will disagree with me. I wonder, however, whether such disagreements are not destined to remain academic. I wonder whether a pure virtue ethic is really an option for human beings. Is it really possible for us to carry on the business of life using (in the final analysis) no moral predicates but those that apply to the characters of moral agents? The record of certain attempts to reform our moral thinking should not encourage those who hope that this will be possible. Since about the middle of the nineteenth century, many thoughtful people have invested considerable energy in the project of eliminating the ancient habit of making moral judgments about people's behavior in favor of judgments about people, judgments expressed in language borrowed from medicine. The invariable result of such projects seems to be that each predicate of the new language of person-evaluation becomes synonymous with some predicate of the old language of action-evaluation—generally "wrong." Thus, we are invited first to say not that it is wicked to murder one's wife, but that the man who does so is "sick," the implication being that such a man should not be punished but should rather receive some sort of medical treatment. If we accept the invitation, however, our use of a new vocabulary does not change the fact that we really do think it is wicked to murder one's wife. Very soon, therefore, we come to mean by calling people "sick" just what we always meant by calling them "wicked." And not long after that, we begin to use the term "sick" for the moral evaluation not only of people but of actions. This process has been completed in our culture, and we now mean by calling an action "sick" what was once meant by calling it "wicked" or "very, very wrong" or "revolting." Indeed, I recently read a story in a newspaper in which a man who had done God knows what awful thing was quoted as saying in his own defense, "What I did wasn't sick—I was mentally ill."

I concede that the therapeutic-reductionist approach to morality is not very much like what people usually have in mind when they speak of the possibility of a virtue ethic. But it is at least logically similar to the typical conception of a virtue ethic in that it proposes to replace the evaluation of actions with the evaluation (or at least the classification) of persons. I predict a similar fate for any

attempt to put a virtue ethic into practice. The evaluative predicates of the theory will first be transferred to actions that are considered typical of the persons to which those predicates apply. And then, owing to the natural tendency of human beings to classify actions as obligatory, permissible, and forbidden (or to employ some system of classification that permits philosophical systemization in terms of these categories), the predicates will detach themselves from their original sense and take on a purely deontological meaning. If there is no free will, then morality as it is ordinarily conceived is an illusion. But I doubt whether it is an illusion from which we shall find it possible to free ourselves.

Notes

1. Utilitarianism, in the strictest sense, is a theory about what is right: that act is right that maximizes utility. But there are many theses that have typically been held by utilitarians that are not a part of utilitarianism in the strictest sense—for example, that utility is to be identified with pleasure. Utilitarianism in the strictest sense would not seem to be any more closely tied to any given theory of the assignment of blame than is any other theory about what is right. I take it that the "typical" utilitarian theory of the assignment of blame is so called because it is the theory of the assignment of blame that has typically been held by utilitarians.
2. It is important to note that the word "fault" in this idiom does not mean "defect." Indeed, in most cases, what is said to be someone's "fault" will not be a property of that person at all. In some cases, moreover, what is said *not* to be someone's fault will be a defect in that person. In certain easily imagined circumstances, one *could* say—but there would seem to be a Gricean rule to the effect that one should not use language that is even superficially paradoxical unless there is some readily apparent reason for doing so—that John has many faults, none of which is his fault.
3. Slote thinks that cases of damage caused by inadvertence more clearly elicit inconsistent moral judgments from people relying on their untutored moral intuitions than do cases of murder. I myself don't see much difference here. I am confident that the merit of the argument I shall present would be unchanged if it were rewritten in terms of an example involving damage caused by inadvertence.
4. By a "pure" virtue ethic I mean a virtue ethic as described in the preceding paragraph: an ethic according to which all moral judgments are reducible to

statements containing no moral predicates but those that apply to a person's character. I believe that sometimes the term "virtue ethic" is used (rather vaguely) in connection with the view that modern moral philosophy has erred in concentrating almost exclusively on obligation and value, thereby neglecting the ancient question, What is the best way for human beings to live? This view is one with which I have considerable sympathy.

Peter van Inwagen
Department of Philosophy
Syracuse University