

## CHAPTER 8

*The Debate over Utilitarianism*

The utilitarian doctrine is that happiness is desirable, and the only thing desirable, as an end; all other things being desirable as means to that end.

JOHN STUART MILL, *UTILITARIANISM* (1861)

Man does not strive after happiness; only the Englishman does that.

FRIEDRICH NIETZSCHE, *TWILIGHT OF THE IDOLS* (1889)

### 8.1. The Resilience of the Theory

*Classical Utilitarianism*—the theory defended by Bentham and Mill—can be summarized in three propositions:

First, actions are to be judged right or wrong solely in virtue of their consequences. Nothing else matters. Right actions are, simply, those that have the best consequences.

Second, in assessing consequences, the only thing that matters is the amount of happiness or unhappiness that is caused. Everything else is irrelevant. Thus right actions are those that produce the greatest balance of happiness over unhappiness.

Third, in calculating the happiness or unhappiness that will be caused, no one's happiness is to be counted as more important than anyone else's. Each person's welfare is equally important. As Mill put it in his *Utilitarianism*,

the happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to

be as strictly impartial as a disinterested and benevolent spectator.

Thus right actions are those that produce the greatest possible balance of happiness over unhappiness, with each person's happiness counted as equally important.

The appeal of this theory to philosophers, economists, and others who theorize about human decision making has been enormous. The theory continues to be widely accepted, even though it has been challenged by a number of apparently devastating arguments. These antiutilitarian arguments are so numerous, and so persuasive, that many have concluded the theory must be abandoned. But the remarkable thing is that so many have *not* abandoned it. Despite the arguments, a great many thinkers refuse to let the theory go. According to these contemporary utilitarians, the antiutilitarian arguments show only that the classical theory needs to be *modified*; they say the basic idea is correct and should be preserved, but recast into a more satisfactory form.

In what follows, we will examine some of these arguments against Utilitarianism, and consider whether the classical version of the theory may be revised satisfactorily to meet them. These arguments are of interest not only for the assessment of Utilitarianism but for their own sakes, as they raise some additional fundamental issues of moral philosophy.

### 8.2. Is Happiness the Only Thing That Matters?

The question *What things are good?* is different from the question *What actions are right?* and Utilitarianism answers the second question by referring back to the first one. Right actions, it says, are the ones that produce the most good. But what is good? The classical utilitarian reply is: one thing, and one thing only, namely happiness. As Mill put it, "The utilitarian doctrine is that happiness is desirable, and the only thing desirable, as an end; all other things being desirable as means to that end."

The idea that happiness is the one ultimate good (and

unhappiness the one ultimate evil) is known as *Hedonism*. Hedonism is a perennially popular theory that goes back at least as far as the ancient Greeks. It has always been an attractive theory because of its beautiful simplicity, and because it expresses the intuitively plausible notion that things are good or bad only on account of the way they make us *feel*. Yet a little reflection reveals serious flaws in the theory. The flaws stand out when we consider examples like these:

1. A promising young pianist's hands are injured in an automobile accident so that she can no longer play. Why is this a bad thing for her? Hedonism would say it is bad because it causes her unhappiness. She will feel frustrated and upset whenever she thinks of what might have been, and *that* is her misfortune. But this way of explaining the misfortune seems to get things the wrong way around. It is not as though, by feeling unhappy, she has made an otherwise neutral situation into a bad one. On the contrary, her unhappiness is a rational response to a situation that *is* unfortunate. She could have had a career as a concert pianist, and now she cannot. *That* is the tragedy. We could not eliminate the tragedy just by getting her to cheer up.

2. You think someone is your friend, but really he ridicules you behind your back. No one ever tells you, so you never know. Is this situation unfortunate for you? Hedonism would have to say no, because you are never caused any unhappiness by the situation. Yet we do feel that there is something bad going on here. You *think* he is your friend, and you are "being made a fool," even though you are not aware of it and so suffer no unhappiness.

Both these examples make the same basic point. We value all sorts of things, including artistic creativity and friendship, for their own sakes. It makes us happy to have them, but only because we *already* think them good. (We do not think them good *because* they make us happy—this is what I meant when I said that Hedonism "gets things the wrong way around.") Therefore we think it a misfortune to lose them, independently of whether or not the loss is accompanied by unhappiness.

In this way, Hedonism misunderstands the nature of happiness. Happiness is not something that is recognized as good

and sought for its own sake, with other things appreciated only as means of bringing it about. Instead, happiness is a response we have to the attainment of things that we recognize *as* goods, independently and in their own right. We think that friendship is a good thing, and so having friends makes us happy. That is very different from first setting out after happiness, then deciding that having friends might make us happy, and then seeking friends as a means to this end.

Today, most philosophers recognize the truth of this. There are not many contemporary hedonists. Those sympathetic to Utilitarianism have therefore sought a way to formulate their view without assuming a hedonistic account of good and evil. Some, such as the English philosopher G. E. Moore (1873–1958), have tried to compile short lists of things to be regarded as good in themselves. Moore suggested that there are three obvious intrinsic goods—pleasure, friendship, and aesthetic enjoyment—and that right actions are those that increase the world's supply of such things. Other utilitarians have tried to bypass the question of how many things are good in themselves, leaving it an open question and saying only that right actions are the ones that have the best results, *however* goodness is measured. Still others try to bypass the question in another way, holding only that we should act so as to maximize the satisfaction of people's *preferences*. It is beyond the scope of this book to discuss the merits or demerits of these varieties of Utilitarianism. I mention them only in order to note that although the hedonistic assumption of the classical utilitarians has largely been rejected, contemporary utilitarians have not found it difficult to carry on. They do so by urging that Hedonism was never a necessary part of the theory in the first place.

### 8.3. Are Consequences All That Matter?

The claim that only consequences matter *is*, however, a necessary part of Utilitarianism. The most fundamental idea underlying the theory is that in order to determine whether an action would be right, we should look at *what will happen as a result of doing it*. If it were to turn out that some *other* matter is

also important in determining rightness, then Utilitarianism would be undermined at its very foundation.

The most serious antiutilitarian arguments attack the theory at just this point: they urge that various other considerations, in addition to utility, are important in determining whether actions are right. We will look briefly at three such arguments.

1. *Justice*. Writing in the academic journal *Inquiry* in 1965, H. J. McCloskey asks us to consider the following case:

Suppose a utilitarian were visiting an area in which there was racial strife, and that, during his visit, a Negro rapes a white woman, and that race riots occur as a result of the crime, white mobs, with the connivance of the police, bashing and killing Negroes, etc. Suppose too that our utilitarian is in the area of the crime when it is committed such that his testimony would bring about the conviction of a particular Negro. If he knows that a quick arrest will stop the riots and lynchings, surely, as a utilitarian, he must conclude that he has a duty to bear false witness in order to bring about the punishment of an innocent person.

This is a fictitious example, but that makes no difference. The argument is only that *if* someone were in this position, then on utilitarian grounds he should bear false witness against the innocent person. This might have some bad consequences—the innocent man might be executed—but there would be enough good consequences to outweigh them: the riots and lynchings would be stopped. The best consequences would be achieved by lying; therefore, according to Utilitarianism, lying is the thing to do. But, the argument continues, it would be wrong to bring about the execution of the innocent man. Therefore, Utilitarianism, which implies it would be right, must be incorrect.

According to the critics of Utilitarianism, this argument illustrates one of the theory's most serious shortcomings: namely, that it is incompatible with the ideal of justice. Justice requires that we treat people fairly, according to their individual needs and merits. The innocent man has done nothing wrong; he did not commit the rape and so he does not deserve to be punished for it. Therefore, punishing him would

be unjust. The example illustrates how the demands of justice and the demands of utility can come into conflict, and so a theory that says utility is the *whole* story cannot be right.

2. *Rights*. Here is a case that is *not* fictitious; it is from the records of the U.S. Court of Appeals, Ninth Circuit (Southern District of California), 1963, in the case of *York v. Story*:

In October, 1958, appellant [Ms. Angelynn York] went to the police department of Chino for the purpose of filing charges in connection with an assault upon her. Appellee Ron Story, an officer of that police department, then acting under color of his authority as such, advised appellant that it was necessary to take photographs of her. Story then took appellant to a room in the police station, locked the door, and directed her to undress, which she did. Story then directed appellant to assume various indecent positions, and photographed her in those positions. These photographs were not made for any lawful purpose.

Appellant objected to undressing. She stated to Story that there was no need to take photographs of her in the nude, or in the positions she was directed to take, because the bruises would not show in any photograph. . . .

Later that month, Story advised appellant that the pictures did not come out and that he had destroyed them. Instead, Story circulated these photographs among the personnel of the Chino police department. In April, 1960, two other officers of that police department, appellee Louis Moreno and defendant Henry Grote, acting under color of their authority as such, and using police photographic equipment located at the police station made additional prints of the photographs taken by Story. Moreno and Grote then circulated these prints among the personnel of the Chino police department. . . .

Ms. York brought suit against these officers and won. Her *legal* rights had clearly been violated. But what of the *morality* of the officers' behavior?

Utilitarianism says that actions are defensible if they produce a favorable balance of happiness over unhappiness. This suggests that we consider the amount of unhappiness caused

to Ms. York and compare it with the amount of pleasure taken in the photographs by Officer Story and his cohorts. It is at least possible that more happiness than unhappiness was caused. In that case, the utilitarian conclusion apparently would be that their actions were morally all right. But this seems to be a perverse way to approach the case. Why should the pleasure afforded Story and his cohorts matter at all? Why should it even count? They had no right to treat Ms. York in that way, and the fact that they enjoyed doing so hardly seems a relevant defense.

To make the point even clearer, consider an (imaginary) related case. Suppose a Peeping Tom spied on Ms. York by peering through her bedroom window, and secretly took pictures of her undressed. Further suppose that he did this without ever being detected and that he used the photographs entirely for his own amusement, without showing them to anyone. Now under these circumstances, it seems clear that the *only* consequence of his action is an increase in his own happiness. No one else, including Ms. York, is caused any unhappiness at all. How, then, could Utilitarianism deny that the Peeping Tom's actions are right? But it is evident to moral common sense that they are not right. Thus, Utilitarianism appears to be an incorrect moral view.

The moral to be drawn from this argument is that Utilitarianism is at odds with the idea that people have *rights* that may not be trampled on merely because one anticipates good results. This is an extremely important notion, which explains why a great many philosophers have rejected Utilitarianism. In the above cases, it is Ms. York's right to privacy that is violated; but it would not be difficult to think of similar cases in which other rights are at issue—the right to freedom of religion, to free speech, or even the right to life itself. It may happen that good purposes are served, from time to time, by ignoring these rights. But we do not think that our rights *should* be set aside so easily. The notion of a personal right is not a utilitarian notion. Quite the reverse: it is a notion that places limits on how an individual may be treated, regardless of the good purposes that might be accomplished.

3. *Backward-Looking Reasons.* Suppose you have promised someone you will do something—say, you promised to meet

him downtown this afternoon. But when the time comes to go, you don't want to do it—you need to do some work and would rather stay home. What should you do? Suppose you judge that the utility of getting your work accomplished slightly outweighs the inconvenience your friend would be caused. Appealing to the utilitarian standard, you might then conclude that it is right to stay home. However, this does not seem correct. The fact that *you promised* imposes an obligation on you that you cannot escape so easily. Of course, if the consequences of not breaking the promise were *great*—if, for example, your mother had just been stricken with a heart attack and you had to rush her to the hospital—you would be justified in breaking it. But a *small* gain in utility cannot overcome the obligation imposed by the fact that you promised. Thus Utilitarianism, which says that consequences are the only things that matter, seems mistaken.

There is an important general lesson to be learned from this argument. Why is Utilitarianism vulnerable to this sort of criticism? It is because the only kinds of considerations that the theory holds relevant to determining the rightness of actions are considerations having to do with the *future*. Because of its exclusive concern with consequences, Utilitarianism has us confine our attention to what *will happen* as a result of our actions. However, we normally think that considerations about the *past* also have some importance. The fact that you promised your friend to meet him is a fact about the past, not the future. Therefore, the general point to be made about Utilitarianism is that it seems to be an inadequate moral theory because it excludes what we might call backward-looking considerations.

Once we understand this point, other examples of backward-looking considerations come easily to mind. The fact that someone did not commit a crime is a good reason why he should not be punished. The fact that someone once did you a favor may be a good reason why you should now do him a favor. The fact that you did something to hurt someone may be a reason why you should now make it up to her. These are all facts about the past that are relevant to determining our obligations. But Utilitarianism makes the past irrelevant, and so it seems deficient for just that reason.

## 8.4. The Defense of Utilitarianism

Taken together, the above arguments form an impressive indictment of Utilitarianism. The theory, which at first seemed so progressive and commonsensical, now seems indefensible: it is at odds with such fundamental moral notions as justice and individual rights, and seems unable to account for the place of backward-looking reasons in justifying conduct. The combined weight of these arguments has prompted many philosophers to abandon the theory altogether.

Many thinkers, however, continue to believe that Utilitarianism, in some form, is true. In reply to the arguments, three general defenses have been offered.

**The First Line of Defense.** The first line of defense is to point out that the examples used in the antiutilitarian arguments are unrealistic and do not describe situations that come up in the real world. Since Utilitarianism is designed as a guide for decision making in the situations we actually face, the fanciful examples are dismissed as irrelevant.

The three antiutilitarian arguments share a common strategy. First a case is described, and then it is noted that from a utilitarian point of view a certain action seems to be required—that is, a certain action would have the best consequences. It is then said that this action is not right. Therefore, it is concluded, the utilitarian conception of rightness cannot be correct.

This strategy succeeds only if we admit that the actions described *really would* have the best consequences. But the utilitarian need not admit this. He can object that, in the real world, bearing false witness does *not* have good consequences. Suppose, in the case described by McCloskey, the “utilitarian” tried to incriminate the innocent man in order to stop the riots. His effort might not succeed; his lie might be found out, and then the situation would be even worse than before. Even if the lie did succeed, the real culprit would remain at large, to commit additional crimes. Moreover, if the guilty party were caught later on, which is always a possibility, the liar would be in deep trouble, and confidence in the criminal justice system would be undermined. The moral is that although

one might *think* that one can bring about the best consequences by such behavior, one can by no means be certain of it. In fact, experience teaches the contrary: utility is not served by framing innocent people. Thus the utilitarian position is *not* at odds with common-sense notions of justice in such cases.

The same goes for the other cases cited in the antiutilitarian arguments. Violating people’s rights, breaking one’s promises, and lying all have bad consequences. Only in philosophers’ imaginations is it otherwise. In the real world, Peeping Toms are caught, just as Officer Story and his cohorts were caught; and their victims suffer. In the real world, when people lie, others are hurt and their own reputations are damaged; and when people break their promises, they lose their friends.

Therefore, far from being incompatible with the idea that we should not violate people’s rights or lie or break our promises, Utilitarianism explains *why* we should not do those things. Moreover, apart from the utilitarian explanation, these duties would remain mysterious and unintelligible. What could be more mysterious than the notion that some actions are right “in themselves,” severed from any notion of a good to be produced by them? Or what could be more unintelligible than the idea that people have “rights” unconnected with any benefits derived from the acknowledgment of those rights? Utilitarianism is not incompatible with common sense; on the contrary, Utilitarianism *is* commonsensical.

**The Second Line of Defense.** The first line of defense contains more bluster than substance. While it can plausibly be maintained that *most* acts of false witness and the like have bad consequences in the real world, it cannot reasonably be asserted that *all* such acts have bad consequences. Surely, in at least some real-life cases, one can bring about good results by doing things that moral common sense condemns. Therefore, in at least some real-life cases Utilitarianism will come into conflict with common sense. Moreover, even if the antiutilitarian arguments had to rely exclusively on fictitious examples, those arguments would nevertheless retain their power; for showing that Utilitarianism has unacceptable consequences in

hypothetical cases is a perfectly valid way of pointing up its theoretical defects. The first line of defense, then, is weak.

The second line of defense admits all this and proposes to save Utilitarianism by giving it a new formulation. In revising a theory to meet criticism, the trick is to identify precisely the feature of the theory that is causing the trouble and to change *that*, leaving the rest of the theory undisturbed as much as possible.

The troublesome aspect of the theory was this: the classical version of Utilitarianism implied that *each individual action* is to be evaluated by reference to its own particular consequences. If on a certain occasion you are tempted to lie, whether it would be wrong is determined by the consequences of *that particular lie*. This, the theory's defenders said, is the point that causes all the trouble; even though we know that *in general* lying has bad consequences, it is obvious that sometimes particular acts of lying can have good consequences.

Therefore, the new version of Utilitarianism modifies the theory so that individual actions will no longer be judged by the Principle of Utility. Instead, *rules* will be established by reference to the principle, and individual acts will then be judged right or wrong by reference to the rules. This new version of the theory is called *Rule-Utilitarianism*, to contrast it with the original theory, now commonly called *Act-Utilitarianism*.

Rule-Utilitarianism has no difficulty coping with the three anti-utilitarian arguments. An act-utilitarian, faced with the situation described by McCloskey, would be tempted to bear false witness against the innocent man because the consequences of *that particular act* would be good. But the rule-utilitarian would not reason in that way. He would first ask, "What *general rules of conduct* tend to promote the greatest happiness?" Suppose we imagine two societies, one in which the rule "Don't bear false witness against the innocent" is faithfully adhered to, and one in which this rule is not followed. In which society are people likely to be better off? Clearly, from the point of view of utility, the first society is preferable. Therefore, the rule against incriminating the innocent should be accepted, and *by appealing to this rule*, the rule-utilitarian

concludes that the person in McCloskey's example should not testify against the innocent man.

Analogous arguments can be used to establish rules against violating people's rights, breaking promises, lying, and so on. We should accept such rules because following them, as a regular practice, promotes the general welfare. But once having appealed to the Principle of Utility to establish the rules, we do not have to invoke the principle again to determine the rightness of particular actions. Individual actions are justified simply by appeal to the already-established rules.

Thus Rule-Utilitarianism cannot be convicted of violating our moral common sense, or of conflicting with ordinary ideas of justice, personal rights, and the rest. In shifting emphasis from the justification of acts to the justification of rules, the theory has been brought into line with our intuitive judgments to a remarkable degree.

**The Third Line of Defense.** Finally, a small group of contemporary utilitarians has had a very different response to the anti-utilitarian arguments. Those arguments point out that the classical theory is at odds with ordinary notions of justice, individual rights, and so on; to this, their response is, essentially, "So what?" In 1961 the Australian philosopher J. J. C. Smart published a monograph entitled *An Outline of a System of Utilitarian Ethics*; reflecting on his position in that book, Smart said:

Admittedly utilitarianism does have consequences which are incompatible with the common moral consciousness, but I tended to take the view "so much the worse for the common moral consciousness." That is, I was inclined to reject the common methodology of testing general ethical principles by seeing how they square with our feelings in particular instances.

Our moral common sense is, after all, not necessarily reliable. It may incorporate various irrational elements, including prejudices absorbed from our parents, our religion, and the general culture. Why should we simply assume that our feelings are always correct? And why should we reject a plausible, rational theory of ethics such as Utilitarianism simply because it

conflicts with those feelings? Perhaps it is the feelings, not the theory, that should be discarded.

In light of this, consider again McCloskey's example of the person tempted to bear false witness. McCloskey argues that it would be wrong to have a man convicted of a crime he did not commit, because it would be unjust. But wait: such a judgment serves *that man's* interests well enough, but what of the *other* innocent people who will be hurt if the rioting and lynchings are allowed to continue? What of them? Surely we might hope that we never have to face a situation like this, for the options are all extremely distasteful. But if we *must* choose between (a) securing the conviction of one innocent person and (b) allowing the deaths of several innocent people, is it so unreasonable to think that the first option, bad as it is, is preferable to the second?

On this way of thinking, Act-Utilitarianism is a perfectly defensible doctrine and does not need to be modified. Rule-Utilitarianism, by contrast, is an unnecessarily watered-down version of the theory, which gives rules a greater importance than they merit. Act-Utilitarianism is, however, recognized to be a radical doctrine which implies that many of our ordinary moral feelings may be mistaken. In this respect, it does what good philosophy always does—it challenges us to rethink matters that we have heretofore taken for granted.

## 8.5. What Is Correct and What Is Incorrect in Utilitarianism

There is a sense in which no moral philosopher can completely reject Utilitarianism. The consequences of one's actions—whether they promote happiness, or cause misery—must be admitted by all to be extremely important. John Stuart Mill once remarked that, insofar as we are benevolent, we must accept the utilitarian standard; and he was surely right. Moreover, the utilitarian emphasis on impartiality must also be a part of any defensible moral theory. The question is whether these are the *only* kinds of considerations an adequate theory must acknowledge. Aren't there *other* considerations that are also important?

If we consult what Smart calls our "common moral con-

sciousness," it seems that there are *many* other considerations that are morally important. (In section 8.3 above, we looked at a few examples.) But I believe the radical act-utilitarians are right to warn us that "common sense" cannot be trusted. Many people once felt that there is an important difference between whites and blacks, so that the interests of whites are somehow more important. Trusting the "common sense" of their day, they might have insisted that an adequate moral theory should accommodate this "fact." Today, no one worth listening to would say such a thing. But who knows how many *other* irrational prejudices are still a part of our moral common sense? At the end of his classic study of race relations, *An American Dilemma* (1944), the Swedish sociologist Gunnar Myrdal reminds us:

There must be still other countless errors of the same sort that no living man can yet detect, because of the fog within which our type of Western culture envelops us. Cultural influences have set up the assumptions about the mind, the body, and the universe with which we begin; pose the questions we ask; influence the facts we seek; determine the interpretation we give these facts; and direct our reaction to these interpretations and conclusions.

The strength of Utilitarianism is that it firmly resists "corruption" by possibly irrational elements. By sticking to the Principle of Utility as the *only* standard for judging right and wrong, it avoids all danger of incorporating into moral theory prejudices, feelings, and "intuitions" that have no rational basis.

The warning should be heeded. "Common sense" can, indeed, mislead us. At the same time, however, there might be at least some nonutilitarian considerations that an adequate theory *should* accept, because there *is* a rational basis for them. Consider, for example, the matter of what people deserve. A person who has worked hard in her job may deserve a promotion more than someone who has loafed, and it would be unjust for the loafer to be promoted first. This is a point that we would expect any fair-minded employer to acknowledge; we would all be indignant if we were passed over for promotion in favor of someone who had not worked as

hard or as well as we. Now utilitarians might agree with this, and say that it can be explained by their theory—they might argue that it promotes the general welfare to encourage hard work by rewarding it. But this does not seem to be an adequate explanation of the importance of desert. The woman who worked harder has a superior claim to the promotion, *not* because it promotes the general welfare for her to get it, but *because she has earned it*. The reason she should be promoted has to do with *her* merits. This does not appear to be the kind of consideration a utilitarian could admit.

Does this way of thinking express a mere prejudice, or does it have a rational basis? I believe it has a rational basis, although it is not one that utilitarians could accept. We ought to recognize individual desert as a reason for treating people in certain ways—for example, as a reason for promoting the woman who has worked harder—because that is the principal way we have of treating individuals as autonomous, responsible beings. If in fact people have the power to choose their own actions, in such a way that they are *responsible* for those actions and what results from them, then acknowledging their deserts is just a way of acknowledging their standing as autonomous individuals. In treating them as they deserve to be treated, we are responding to the way they have freely chosen to behave. Thus in some instances we will not treat everyone alike, because people are not just members of an undifferentiated crowd. Instead, they are individuals who, by their own choices, show themselves to deserve different kinds of responses.

I will say more about what people deserve, and why this is important for ethics, in Chapters 10 and 13. Here I will draw only this conclusion about Utilitarianism: although it emphasizes points that any adequate moral theory must acknowledge, Utilitarianism is not itself a fully adequate theory because there is at least one important matter—individual desert—that escapes its net.