7

Utilitarianism

Suppose you are on an island with a dying millionaire. With his final words, he begs you for one final favor: "I've dedicated my whole life to baseball and for fifty years have gotten endless pleasure rooting for the New York Yankees. Now that I am dying, I want to give all my assets, $5 million, to the Yankees." Pointing to a box containing money in large bills, he continues: "Would you take this money back to New York and give it to the Yankees' owner so that he can buy better players?" You agree to carry out his wish, at which point a huge smile of relief and gratitude breaks out on his face as he expires in your arms. After traveling to New York, you see a newspaper advertisement placed by your favorite charity, World Hunger Relief Organization (whose integrity you do not doubt), pleading for $5 million to be used to save 100,000 people dying of starvation in Africa. Not only will the $5 million save their lives, but it will also purchase equipment and the kinds of fertilizers necessary to build a sustainable economy. You decide to reconsider your promise to the dying Yankee fan, in light of this advertisement.

What is the right thing to do in this case? Consider some traditional moral principles and see if they help us come to a decision. One principle often given to guide action is "Let your conscience be your guide." I recall this principle with fondness, for it was the one my father taught me at an early age, and it still echoes in my mind. But does it help here? No, since conscience is primarily a function of upbringing. People's consciences speak to them in different ways according to how they were brought up. Depending on upbringing, some people feel no qualms about committing violent acts, whereas others feel the torments of conscience over stepping on a gnat. Suppose your conscience tells you to give the money to the Yankees and my conscience tells me to give the

money to the World Hunger Relief Organization. How can we even discuss the matter? If conscience is the end of it, we're left mute.

Another principle urged on us is "Do whatever is most loving"; Jesus in particular set forth the principle "Love your neighbor as yourself." Love is surely a wonderful value. It is a more wholesome attitude than hate, and we should overcome feelings of hate if only for our own psychological health. But is love enough to guide our actions when there is a conflict of interest? "Love is blind," it has been said, "but reason, like marriage, is an eye-opener." Whom should I love in the case of the disbursement of the millionaire's money—the millionaire or the starving people? It's not clear how love alone will settle anything. In fact, it is not obvious that we must always do what is most loving. Should we always treat our enemies in loving ways? Or is it morally permissible to feel hate for those who have purposely and unjustly harmed us, our loved ones, or other innocent people? Should the survivors of Nazi concentration camps love Adolph Hitler? Love alone does not solve difficult moral issues.

A third principle often given to guide our moral actions is the Golden Rule: "Do to others as you would have them do to you." This, too, is a noble rule of thumb, one that works in simple, commonsense situations. But it has problems. First, it cannot be taken literally. Suppose I love to hear loud heavy-metal music. Since I would want you to play it loudly for me, I reason that I should play it loudly for you—even though I know that you hate the stuff. Thus, the rule must be modified: "Do to others as you would have them do to you if you were in their shoes." However, this still has problems. If I were the assassin of Robert Kennedy, I'd want to be released from the penitentiary; but it is not clear that he should be released. If I put myself in the place of a sex-starved individual, I might want to have sex with the next available person; but it's not obvious that I (or anyone else) must comply with that wish. Likewise, the Golden Rule doesn't tell me to whom to give the millionaire's money.

Conscience, love, and the Golden Rule are all worthy rules of thumb to help us through life. They work for most of us, most of the time, in ordinary moral situations. But, in more complicated cases, especially when there are legitimate conflicts of interests, they are limited.

A more promising strategy for solving dilemmas is that of following definite moral rules. Suppose you decided to give the millionaire's money to the Yankees to keep your promise or because to do otherwise would be stealing. The principle you followed would be "Always keep your promise." Principles are important in life. All learning involves understanding a set of rules; as R. M. Hare says, "Without principles we could not learn anything whatever from our elders.... Every generation would have to start from scratch and teach itself."[1] If you

decided to act on the principle of keeping promises, then you adhered to a type of moral theory called **deontology**. In Chapter 1, we saw that deontological systems maintain that the center of value is the act or kind of act; certain features in the act itself have intrinsic value. For example, a deontologist would see something intrinsically wrong in the very act of lying.

If, on the other hand, you decided to give the money to the World Hunger Relief Organization to save an enormous number of lives and restore economic solvency to the region, you sided with a type of theory called **teleological ethics**. Sometimes, it is referred to as *consequentialist ethics*. We also saw in Chapter 1 that the center of value here is the outcome or consequences of the act. For example, a teleologist would judge whether lying was morally right or wrong by the consequences it produced.

We have already examined one type of teleological ethics: *ethical egoism*, the view that the act that produces the most amount of good for the agent is the right act. Egoism is teleological ethics narrowed to the agent himself or herself. In this chapter, we will consider the dominant version of teleological ethics—*utilitarianism*. Unlike ethical egoism, utilitarianism is a universal teleological system. It calls for the maximization of goodness in society—that is, the greatest goodness for the greatest number—and not merely the good of the agent.

CLASSIC UTILITARIANISM

In our normal lives we use utilitarian reasoning all the time; I might give money to charity when seeing that it would do more good for needy people than it would for me. In time of war, I might join the military and risk dying because I see that society's needs at that time are greater than my own. As a formal ethical theory, the seeds of utilitarianism were sewn by the ancient Greek philosopher Epicurus (342–270 BCE), who stated that "pleasure is the goal that nature has ordained for us; it is also the standard by which we judge everything good." According to this view, rightness and wrongness are determined by pleasure or pain that something produces. Epicurus's theory focused largely on the individual's personal experience of pleasure and pain, and to that extent he advocated a version of ethical egoism. Nevertheless, Epicurus inspired a series of eighteenth-century philosophers who emphasized the notion of general happiness—that is, the pleasing consequences of actions that impact others and not just the individual. Francis Hutcheson (1694–1746) stated that "that action is best, which procures the greatest happiness for the greatest numbers." David Hume (1711–1776) introduced the term *utility* to describe the pleasing consequences of actions as they impact people.

The classical expressions of utilitarianism, though, appear in the writings of two English philosophers and social reformers Jeremy Bentham (1748–1832) and

John Stuart Mill (1806–1873). Their approach to morality was nonreligious and they tried to reform society by rejecting unfounded rules of morality and law.

Jeremy Bentham

There are two main features of utilitarianism, both of which Bentham articulated: the consequentialist principle (or its teleological aspect) and the utility principle (or its hedonic aspect). The *consequentialist principle* states that the rightness or wrongness of an act is determined by the goodness or badness of the results that follow from it. It is the end, not the means, that counts; the end justifies the means. The *utility*, or *hedonist, principle* states that the only thing that is good in itself is some specific type of state (for example, pleasure, happiness, welfare). Hedonistic utilitarianism views pleasure as the sole good and pain as the only evil. To quote Bentham, "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as what we shall do."[2] An act is right if it either brings about more pleasure than pain or prevents pain, and an act is wrong if it either brings about more pain than pleasure or prevents pleasure from occurring.

Bentham invented a scheme for measuring pleasure and pain that he called the **hedonic calculus**. The quantitative score for any pleasure or pain experience is obtained by summing the seven aspects of a pleasurable or painful experience: its intensity, duration, certainty, nearness, fruitfulness, purity, and extent. Adding up the amounts of pleasure and pain for each possible act and then comparing the scores would enable us to decide which act to perform. With regard to our example of deciding between giving the dying man's money to the Yankees or to the African famine victims, we would add up the likely pleasures to all involved, for all seven qualities. If we found that giving the money to the famine victims would cause at least 3 million *hedons* (units of happiness) but that giving the money to the Yankees would cause less than 1,000 hedons, we would have an obligation to give the money to the famine victims.

There is something appealing about Bentham's utilitarianism. It is simple in that there is only one principle to apply: Maximize pleasure and minimize suffering. It is commonsensical in that we think that morality really is about reducing suffering and promoting benevolence. It is scientific: Simply make quantitative measurements and apply the principle impartially, giving no special treatment to ourselves or to anyone else because of race, gender, personal relationship, or religion.

However, Bentham's philosophy may be too simplistic in one way and too complicated in another. It may be too simplistic in that there are values other than pleasure (as we saw in Chapter 6), and it seems too complicated in its artificial hedonic calculus. The calculus is burdened with too many variables and has problems assigning scores to the variables. For instance, what score do we give a cool drink on a hot day or a warm shower on a cool day? How do we compare a 5-year-old's delight over a new toy with a 30-year-old's delight with a new lover? Can we take your second car from you and give it to Beggar Bob, who does not own a car and would enjoy it more than you? And if it is simply the overall benefits

of pleasure that we are measuring, then if Jack or Jill would be "happier" in the Pleasure Machine or the Happiness Machine or on drugs than in the real world, would we not have an obligation to ensure that these conditions become reality? Because of such considerations, Bentham's version of utilitarianism was, even in his own day, referred to as the "pig philosophy" because a pig enjoying his life would constitute a higher moral state than a slightly dissatisfied Socrates.

John Stuart Mill

It was to meet these sorts of objections and save utilitarianism from the charge of being a pig philosophy that Bentham's successor, John Stuart Mill, sought to distinguish happiness from mere sensual pleasure. His version of the theory is often called **eudaimonistic utilitarianism** (from the Greek *eudaimonia*, meaning "happiness"). He defines happiness in terms of certain types of higher-order pleasures or satisfactions such as intellectual, aesthetic, and social enjoyments, as well as in terms of minimal suffering. That is, there are two types of pleasures. The lower, or elementary, include eating, drinking, sexuality, resting, and sensuous titillation. The higher include high culture, scientific knowledge, intellectuality, and creativity. Although the lower pleasures are more intensely gratifying, they also lead to pain when overindulged in. The higher pleasures tend to be more long term, continuous, and gradual.

Mill argued that the higher, or more refined, pleasures are superior to the lower ones: "A being of higher faculties requires more to make him happy, is capable probably of more acute suffering, and certainly accessible to it at more points, than one of an inferior type," but still he is qualitatively better off than the person without these higher faculties. "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied."[3] Humans are the kind of creatures who require more to be truly happy. They want the lower pleasures, but they also want deep friendship, intellectual ability, culture, the ability to create and appreciate art, knowledge, and wisdom.

But one may object, "How do we know that it really is better to have these higher pleasures?" Here, Mill imagines a panel of experts and says that of those who have had a wide experience of pleasures of both kinds almost all give a decided preference to the higher type. Because Mill was an *empiricist*—one who believed that all knowledge and justified belief was based on experience—he relied on the combined consensus of human history. By this view, people who experience both rock music and classical music will, if they appreciate both, prefer Bach and Beethoven to Metallica. That is, we generally move up from appreciating simple things (for example, nursery rhymes) to more complex and intricate things (for example, poetry that requires great talent) rather than the other way around.

Mill has been criticized for not giving a better reply—for being an elitist and for unduly favoring the intellectual over the sensual. But he has a point. Don't we generally agree, if we have experienced both the lower and the higher types of pleasure, that even though a full life would include both, a life with only the former is inadequate for human beings? Isn't it better to be Socrates dissatisfied than the pig satisfied—and better still to be Socrates satisfied?

The point is not merely that humans wouldn't be satisfied with what satisfies a pig but that somehow the quality of the higher pleasures is *better*. But what does it mean to speak of better pleasure? The formula he comes up with is this:

> Happiness … [is] not a life of rapture; but moments of such, in an existence made up of few and transitory pains, many and various pleasures, with a decided predominance of the active over the passive, and having as the foundation of the whole, not to expect more from life than it is capable of bestowing.[4]

Mill is clearly pushing the boundaries of the concept of "pleasure" by emphasizing higher qualities such as knowledge, intelligence, freedom, friendship, love, and health. In fact, one might even say that his litmus test for happiness really has little to do with actual pleasure and more to do with a nonhedonic-cultivated state of mind.

ACT- AND RULE-UTILITARIANISM

There are two classical types of utilitarianism: act- and rule-utilitarianism. In applying the principle of utility, act-utilitarians, such as Bentham, say that ideally we ought to apply the principle to all of the alternatives open to us at any given moment. We may define act-utilitarianism in this way:

> **Act-utilitarianism**: An act is right if and only if it results in as much good as any available alternative.

One practical problem with act-utilitarianism is that we cannot do the necessary calculations to determine which act is the correct one in each case, for often we must act spontaneously and quickly. So rules of thumb are of practical importance—for example, "In general, don't lie," and "Generally, keep your promises." However, the right act is still that alternative that results in the most utility.

A second problem with act-utilitarianism is that it seems to fly in the face of fundamental intuitions about minimally correct behavior. Consider Richard Brandt's criticism of act-utilitarianism:

> It implies that if you have employed a boy to mow your lawn and he has finished the job and asks for his pay, you should pay him what you promised only if you cannot find a better use for your money. It implies that when you bring home your monthly paycheck you should use it to support your family and yourself only if it cannot be used more effectively to supply the needs of others.[5]

The alternative to act-utilitarianism is a view called rule-utilitarianism—elements of which we find in Mill's theory. Most generally, the position is this:

> **Rule-utilitarianism**: An act is right if and only if it is required by a rule that is itself a member of a set of rules whose acceptance would lead to greater utility for society than any available alternative.

Human beings are rule-following creatures. We learn by adhering to the rules of a given subject, whether it is speaking a language, driving a car, dancing, writing an essay, rock climbing, or cooking. We want to have a set of action-guiding rules by which to live. The act-utilitarian rule, to do the act that maximizes utility, is too tedious for most purposes. Often, we don't have time to decide whether lying will produce more utility than truth telling, so we need a broad rule prescribing truthfulness that passes the test of rational scrutiny. Rule-utilitarianism asserts that the best chance of maximizing utility is by following the *set of rules* most likely to give us our desired results. Because morality is a social and public institution, we need to coordinate our actions with others so that we can have reliable expectations about other people's behavior.

For the most sophisticated versions of rule-utilitarianism, three levels of rules will guide actions. On the lowest level is a set of utility-maximizing rules of thumb, such as "Don't lie" and "Don't cause harm," that should always be followed unless there is a conflict between them. If these first-order rules conflict, then a second-order set of conflict-resolving rules should be consulted, such as "It's more important to avoid causing serious harm than to tell the truth." At the top of the hierarchy is a third-order rule sometimes called the *remainder rule*, which is the principle of act-utilitarianism: When no other rule applies, simply do what your best judgment deems to be the act that will maximize utility.

An illustration of this is the following: Suppose you promised to meet your teacher at 3 p.m. in his office. On your way there, you come upon an accident victim stranded by the wayside who desperately needs help. The two first-order rules in this situation are "Keep your promises" and "Help those in need when you are not seriously inconvenienced in doing so." It does not take you long to decide to break the appointment with your teacher because it seems obvious in this case that the rule to help others overrides the rule to keep promises. There is a second-order rule prescribing that the first-order rule of helping people in need when you are not seriously inconvenienced in doing so overrides the rule to keep promises. However, there may be some situation where no obvious rule of thumb applies. Say you have $50 that you don't really need now. How should you use this money? Put it into your savings account? Give it to your favorite charity? Use it to throw a party? Not only is there no clear first-order rule to guide you, but there is no second-order rule to resolve conflicts between first-order rules. Here and only here, on the third level, the general act-utility principle applies without any other primary rule; that is, do what in your best judgment will do the most good.

Debates between act- and rule-utilitarians continue today. Kai Nielsen, a staunch act-utilitarian, argues that no rules are sacred; differing situations call forth different actions, and potentially any rule could be overridden. He thus criticizes what he calls *moral conservatism*, which is any normative ethical theory that maintains that there is a privileged moral principle, or cluster of moral principles, prescribing determinate actions that it would always be wrong not to act in accordance with no matter what the consequences.

Nielsen argues further that we are responsible for the consequences of not only the actions that we perform but also the nonactions that we fail to perform. He calls

this "negative responsibility." To illustrate, suppose you are the driver of a trolley car and suddenly discover that your brakes have failed. You are just about to run over five workers on the track ahead of you. However, if you act quickly, you can turn the trolley onto a sidetrack where only one man is working. What should you do? One who makes a strong distinction between *allowing* versus *doing* evil would argue that you should do nothing and merely allow the trolley to kill the five workers. But one who denies that this is an absolute distinction would prescribe that you do something positive to minimize evil. Negative responsibility means that you are going to be responsible for someone's death in either case. Doing the right thing, the utilitarian urges, means minimizing the amount of evil. So you should actively cause the one death to save the other five lives.[6] Critics of utilitarianism contend either that negative responsibility is not a strict duty or that it can be worked into other systems besides utilitarianism.

The Strengths of Utilitarianism

Utilitarianism has three positive features. The first attraction or strength is that it is a single principle, an absolute system with a potential answer for every situation: Do what will promote the most utility. It's good to have a simple, action-guiding principle that is applicable to every occasion—even if it may be difficult to apply (life's not simple).

Its second strength is that utilitarianism seems to get to the substance of morality. It is not merely a formal system that simply sets forth broad guidelines for choosing principles but offers no principles—such as the guideline "Do whatever you can universalize." Rather it has a material core: We should promote human (and possibly animal) flourishing and reduce suffering. The first virtue gives us a clear decision procedure in arriving at our answer about what to do. The second virtue appeals to our sense that morality is made for people and that morality is not so much about rules as about helping people and alleviating the suffering in the world.

As such, utilitarianism seems commonsensical. For instance, it gives us clear and reasonable guidance in dealing with the Kitty Genovese case discussed in Chapter 1: We should call the police or do what is necessary to help her, as long as helping her does not create more disutility than leaving her alone. And, in the case of deciding what to do with the dead millionaire's $2 million, something in us says that it is absurd to keep a promise to a dead person when it means allowing hundreds of thousands of famine victims to die. Far more good can be accomplished by helping the needy than by giving the money to the Yankees.

A third strength of utilitarianism is that it is particularly well suited to address the **problem of posterity**—namely, why we should preserve scarce natural resources for the betterment of future generations of humans that do not yet exist. Expressed rhetorically, the question is "Why should I care about posterity; what has posterity ever done for me?" In Chapter 6, we saw that the theory of ethical egoism failed to give us an adequate answer to this problem. That is, the egoist gains nothing by preserving natural resources for future generations that do not yet exist and thus can give no benefit to the egoist. However, utilitarians

have one overriding duty: to maximize general happiness. As long as the quality of life of future people promises to be positive, we have an obligation to continue human existence, to produce human beings, and to take whatever actions are necessary to ensure that their quality of life is not only positive but high.

It does not matter that we cannot identify these future people. We may look upon them as mere abstract placeholders for utility and aim at maximizing utility. Derek Parfit explains this using this utilitarian principle: "It is bad if those who live are worse off than those who might have lived." He illustrates his principle this way. Suppose our generation has the choice between two energy policies: the "Safe Energy Policy" and the "Risky Energy Policy."[7] The Risky Policy promises to be safe for us but is likely to create serious problems for a future generation, say, 200 years from now. The Safe Policy won't be as beneficial to us but promises to be stable and safe for posterity—those living 200 years from now and beyond. We must choose and we are responsible for the choice that we make. If we choose the Risky Policy, we impose harms on our descendants, even if they don't now exist. In a sense, we are responsible for the people who will live because our policy decisions will generate different causal chains, resulting in different people being born. But more important, we are responsible for their quality of life because we could have caused human lives to have been better off than they are.

What are our obligations to future people? If utilitarians are correct, we have an obligation to leave posterity to as good a world as we can. This would mean radically simplifying our lifestyles so that we use no more resources than are necessary, keeping as much top soil intact as possible, protecting endangered species, reducing our carbon dioxide emissions, preserving the wilderness, and minimizing our overall deleterious impact on the environment in general while using technology wisely.

CRITICISM OF UTILITARIANISM

Utilitarianism has been around for several centuries, but so too have been its critics, and we need to address a series of standard objections to utilitarianism before we can give it a "philosophically clean bill of health."

Problems with Formulating Utilitarianism

The first set of problems occurs in the very formulation of utilitarianism: "The greatest happiness for the greatest number." Notice that we have two "greatest" things in this formula: "happiness" and "number." Whenever we have two variables, we invite problems of determining which of the variables to rank first when they seem to conflict. To see this point, consider the following example: I am offering a $1,000 prize to the person who runs the longest distance in the shortest amount of time. Three people participate: Joe runs 5 miles in 31 minutes, John runs 7 miles in 50 minutes, and Jack runs 1 mile in 6 minutes. Who should get the prize? John has fulfilled one part of the requirement (run the

longest distance), but Jack has fulfilled the other requirement (run the shortest amount of time).

This is precisely the problem with utilitarianism. On the one hand, we might concern ourselves with spreading happiness around so that the greatest number obtain it (in which case, we should get busy and procreate a larger population). On the other hand, we might be concerned that the greatest possible amount of happiness obtains in society (in which case, we might be tempted to allow some people to become far happier than others, as long as their increase offsets the losers' diminished happiness). So should we worry more about total happiness or about highest average?

Utilitarians also need to be clear about specifically whose happiness we are talking about: all beings that experience pleasure and pain, or all human beings, or all rational beings. One criterion might exclude mentally deficient human beings, and another might include animals. Finally, utilitarians need to indicate how we measure happiness and make interpersonal comparisons between the happiness of different people. We've seen Mill's efforts to address this problem with his notion of higher pleasures; we've also seen the additional complications that his solution creates.

None of these problems defeat utilitarianism as a workable theory, but they do place a heavy burden on utilitarians to clarify the objectives of their theory.

The Comparative Consequences Objection

Another crucial problem with utilitarianism is that it seems to require a superhuman ability to look into the future and survey a mind-boggling array of consequences of actions. Of course, we normally do not know the long-term consequences of our actions because life is too complex and the consequences go on into the indefinite future. One action causes one state of affairs, which in turn causes another state of affairs, indefinitely, so that calculation becomes impossible. Recall the nursery rhyme:

> For want of a nail, the shoe was lost;
> For want of a shoe, the horse was lost;
> For want of a horse, the rider was lost;
> For want of a rider, the battle was lost;
> For want of a battle, the kingdom was lost;
> And all for the want of a horseshoe nail.

Poor, unfortunate blacksmith; what utilitarian guilt he must bear all the rest of his days!

But it is ridiculous to blame the loss of one's kingdom on the poor, unsuccessful blacksmith, and utilitarians are not so foolish as to hold him responsible for the bad situation. Instead, following C. I. Lewis, utilitarians distinguish two kinds of consequences: (1) actual consequences of an act and (2) consequences that could reasonably have been expected to occur.[8] Based on these two kinds of consequences, there are two corresponding right actions. An act is *absolutely* right if it has the best actual consequences (as per consequence 1). An act is