

## I. Utilitarian Ethics

### *Utilitarianism as a Consequentialist Moral Theory*

Utilitarianism is the leading theory in a general category of theories known as consequentialism. The individual theories in the category of consequentialism will differ according to how the consequences are to be examined. More specifically, any consequentialist theory will consist of two main parts, the good and the right. There are various possibilities for what might count as the good and for what might count as the right, and the nature of any specific consequentialist theory will be a function of which possibility is chosen for each.

First, the good specifies what goal(s) is morally worthy of pursuit. In other words, it must be made clear what sorts of consequences are desirable and what sorts are undesirable, and this can be done only when the goal is clear. Actions contributing to the achievement of that goal are then said to be morally desirable, while actions inhibiting its achievement are said to be undesirable. This goal is referred to by philosophers as the good. Any consequentialist theory must therefore provide a theory of the good, which is an account of the goal to be pursued. Some examples include a welfarist theory of the good, according to which individual well-being is the relevant goal worthy of pursuit, and a liberty-based theory of the good, according to which freedom is the morally relevant goal. One influential philosopher, G. E. Moore (1873–1958), even argued that the nature of the good is primitive and indefinable, and for this reason cannot be specified in language but can only be grasped intuitively. Classical utilitarianism asserts a broad notion of happiness as the good.

Second, the right is a more specific indication of what is to be done with the good. The most common theory of the right is maximization. Simply, this stipulates that the good, whatever it happens to be, should be maximized; the more of it the better. On this approach, an action is morally right when it contributes to the achievement of the good, and the more it contributes to the good the more morally right it is. Conversely, an action is morally wrong when it hinders the achievement of the good, and the greater the hindrance created, the more morally wrong the action. Other theories of the right suggest that the good need not be maximized. For example, a “satisficing” theory of the right stipulates that the good must be achieved only to a satisfactory degree. Taking happiness as the good, it might be the case that a company’s benefit package is fairly comprehensive and thus generates a reasonable (satisfactory) amount of happiness. With some additions, however, the benefit package would make the employees much more happy, so much so that any unhappiness incurred by the company’s managers (who would have to pay out the extra benefits) would be more than offset. Thus, the result would be an increase in overall happiness. The maximizing approach would require the additions to the benefit package; indeed, providing the merely satisfactory version would be ethically wrong, as it would fail to maximize happiness. The satisficing approach would not require the additions, since the previous version produced a satisfactory amount of happiness. A third possible sort of theory would incorporate distributive concerns, requiring that the good be allocated to people in certain ways; some philosophers have suggested an equal distribution, while others have argued that the right requires that every individual enjoy at least a minimally acceptable level of the good. Classical utilitarianism adopts the maximizing theory of the good, claiming that this is the only approach that makes sense. Whatever the good happens to be, it only seems logical that more of that good will produce even better consequences.

### Basic Aspects of Utilitarianism

(When the happiness theory of the good and the maximizing theory of the right are combined, the result is a basic definition of utilitarianism. Specifically, utilitarianism requires the maximization of overall happiness. *Maximization* refers to the theory of the right and indicates that the more happiness there is, the better. *Overall* indicates that the total happiness is what is relevant. There is sometimes a tendency to think in terms of maximizing “my” happiness or “their” happiness—maximizing the happiness of this group (such as a company) or that group (such as a local community). This is not the correct way to apply utilitarianism. Rather, the effects of an action on the happiness of *all* people must be taken into consideration. Often, large segments of the population will be unaffected by an action that might be taken, but it would be misleading (and even wrong) to say that unaffected people are not being considered in the ethical analysis. The accurate description, rather, is that all people are being considered, but it is efficient to focus only on those people for whom the act in question would produce some net change in happiness.)

*Happiness* is perhaps the trickiest term in the basic definition provided here. Some references to utilitarianism point to utility as the good rather than happiness, but the two are roughly synonymous; the idea is to specify a good that is general enough to capture a variety of other, more specific possible terms. However, this generality must be balanced with some specifics. Jeremy Bentham (1748–1832) is the philosopher most often cited as providing the first complete account of utilitarianism, including an account of happiness. According to Bentham, *happiness* is a hedonistic notion, founded on the natural state of humankind, and is a function of achieving pleasure and avoiding pain. He even worked out an “*hedonic calculus*” that assigned relative values, positive and negative, to various acts that would produce pleasure and pain respectively. The ethically correct course of action, the one that maximizes overall happiness, could then be determined by filling in the various values and comparing the overall sums generated by each possible alternative. Bentham’s version of utilitarianism thus requires measuring specific quantities of happiness, and while this might seem a strange aspect of the theory, utilitarians argue that some such approach must be adopted, even if the numbers are a bit imprecise. Examples of how this might work will be seen momentarily.

The *hedonic calculus*, however, was the target of some criticism, even in Bentham’s own day and even by other philosophers sympathetic to the basic utilitarian idea. In addition to being difficult to manage, the calculus at times led to fairly strange conclusions, including the moral obligation to prioritize the well-being of animals over that of human beings in various situations. It was this result in particular that led some to condemn Bentham’s utilitarianism as a “*pig philosophy*.” John Stuart Mill (1806–1873) attempted to salvage the theory by refuting the quantitative approach of Bentham and relying instead on a more general conception of happiness. Mill distinguished “higher” pleasures from “lower” pleasures; the former pertain to the ability to reason and thus can be enjoyed only by rational beings, while the latter pertain to more basic functions and can be enjoyed by beings that are not rational. Examples of higher pleasures include intellectuality and creativity; examples of lower pleasures include eating, drinking, and sex. The final aspect of Mill’s approach is the assertion that the higher pleasures are superior to the lower ones—that given the choice, a person would prefer to possess reason and creativity regardless of how unsatisfactory his life might seem to him. This is the basis for Mill’s well-known quotation that, “It is better to be a human being dissatisfied than a pig satisfied.” So in addition to his allegedly saving utilitarianism from being a “*pig philosophy*,”

Mill also maintained that the messiness associated with quantifying happiness could be avoided with the “higher-lower pleasure” restructuring he proposed.

### *Act-Utilitarianism*

The discussion of utilitarianism to this point has in essence made use of one version of the theory, act-utilitarianism, which will now be explained in more detail. The other principal version, rule-utilitarianism, will be discussed later.

The definition provided earlier, which called for the maximization of overall happiness, can be used in connection with act-utilitarianism. More formally, act-utilitarianism is the theory stipulating that the morally right act is the one that produces at least as much overall happiness in the circumstances as any alternative act. This means that when deciding which act would be ethically right, a person must investigate the reasonably foreseeable consequences of the different possible acts she could perform; the act that would produce the most overall happiness is morally right and must be carried out.

A simple example, with just two parties involved and just two alternative courses of action, can be used to illustrate how this would work. Consider a request made by your aunt in the local retirement home that you visit her one evening. She can be a difficult person and you do not particularly like visiting, though you know she is lonely and your visits do her a world of good. If you agree to her request and visit her, she will be very happy, but you will miss your favorite television show, causing you to experience a certain amount of unhappiness. If you decide not to visit her, she will be extremely disappointed, but you will be able to watch your television program, gaining some happiness (though a twinge of guilt takes away just a bit of that happiness). Act-utilitarianism requires this sort of thinking—analyzing the effects on the happiness of the individuals involved for each alternative course of action. The morally right action is the one producing the most overall happiness. The following chart captures these various effects:

	Visit	Do Not Visit
You	-10	+8
Aunt	+15	-20
Total	+5	-12

Notice that the amount of happiness gained or lost is quantified. It may therefore be thought that Bentham’s version of act-utilitarianism, with its hedonic calculus of “units” of happiness, has been applied here. This is not necessarily the case, however. Despite Mill’s attempt to make use of a more general notion of happiness, some indication of the degree to which people are made happy or unhappy must be provided; unless the strength of people’s preferences or aversions are taken into account, the outcome will not provide a true indication of the overall effect on happiness but will instead be something like a vote tally. The “vote” in this case would be split, you voting not to visit and your aunt voting that you do visit. Only when the strengths of preferences are considered will the more precise overall effects on happiness become evident. As indicated in the chart, the world would be happier (by 5 “units”) if you were to visit your aunt and it would be less happy (by 12 “units”) if you were to stay home and watch television. This follows from your aunt having more happiness at stake. Assuming these estimates are correct and assuming there are no other parties whose happiness would be affected by your visiting or not visiting, the act-utilitarian would say that it is your moral duty to visit your aunt. (One would also have to think about whether there might be any alternatives other than the two that are considered.) This example, while oversimplified, is consistent with both

Bentham's approach (which makes use of specific numerical quantities based on pleasure and pain) and Mill's approach (which, despite the generality and the higher-lower pleasure distinction, must allow for degrees of preferences to be taken into account).

In line with the formal definition of act-utilitarianism provided here, it should also be pointed out that in the case where there is a "tie" in that two alternative courses of action would produce the most happiness, either of the two acts may be performed. This sort of possibility is covered by the specification that the right act would produce "at least" as much overall happiness as any alternative act. If, for example, the total effects on happiness would be identical for Act X and Act Y (if, say, each would produce a net gain of 5 units of happiness), then either of those two actions would be morally permissible. To say that an act is morally *permissible*, recall, is to say that it is allowed but not required; the performance of the act is not morally wrong, but neither is the failure to perform it. However, performing one of those two acts, instead of Act Z (that would, say, produce an overall gain of only 2 units of happiness), would be morally required.

Act-utilitarianism has been a popular moral theory for some time, and philosophers who find it to be plausible point to several considerations in its favor. First, it assesses the morality of actions in terms of the consequences of those actions, and utilitarians suggest that this is commonsensical, that what we mean by the morality of actions is just the consequences that those actions might bring about. In what other ways, they ask, could morality be assessed except by appealing to the consequences? Second, it is relatively simple and (all things considered) easy to apply; the principle "maximize overall happiness" is straightforward. This is important, according to utilitarians, because in areas of applied ethics (such as professional ethics), it is desirable to have a basic moral theory that can be applied to real-life cases without undue difficulty. Third, act-utilitarianism possesses the characteristics described in the Introduction, that are important for a viable moral theory. (In addition to being prescriptive, for example, it is impartial; every individual is treated equally, in that no one person's happiness is weighted more heavily than any other person's. Differences exist among individuals only in the degree of happiness (or unhappiness) that would be brought about, but act-utilitarianism treats these differences in a consistent, impartial manner. In this way, all parties are treated the same.)

**Criticisms of Act-Utilitarianism** This moral theory has also been the subject of a variety of objections. Several potential problems have been described by critics—problems which, if coherent, could raise serious doubts about the plausibility of act-utilitarianism as a workable guideline for professional ethics, or for ethics in any context.

First, it has been claimed that act-utilitarianism is not practical, as it requires decision-makers to measure happiness, and this is not feasible. At some point in the ethical analysis, some sort of measurement will be required, as was made clear in the example of you visiting your aunt; some assessment of the degree of happiness (or unhappiness) created is necessary. Part of the problem concerns the definition of *happiness* and the pitfalls of determining how very different sorts of happiness can be measured and compared. This problem was cited by the philosopher Alasdair MacIntyre, who claimed that the enjoyment derived from a swim and the enjoyment derived from a cool drink, for example, are very different, and attempts to combine the two under any coherent definition of *happiness* cannot be successful. Another part of the problem is that act-utilitarianism requires insights into the internal mental states of others. It requires the decision-maker to grasp the type and degree of happiness that would be produced in others, and this is of course difficult if not impossible. At the least it requires reliable estimations of the measurements of others' happiness, and this will be an extremely imprecise exercise.

Second, because it is concerned strictly with the outcomes of actions that have not been (and may never be) performed, act-utilitarianism has been criticized for requiring decision-makers to predict the future. Because predicting the future is always an uncertain business, act-utilitarianism requires decision-makers to do what they are in fact not able to do, and this makes the theory impractical. This second criticism and the first one have in common the general problem of lack of information. If the various calculations can indeed be assembled—calculations that accurately capture the happiness of all parties involved and that accurately predict what the effects on happiness will be—then the mechanical process of adding numbers and comparing sums of happiness will of course be simple and straightforward. The first two criticisms, however, suggest that there are significant problems with the gathering of information that would be relevant to the calculations.

Third, it has been claimed that act-utilitarianism, when taken to its logical extreme, requires too much of people. In fact, it follows from the basic tenets of the theory that we are morally required to engage in never-ending action. This becomes clear, according to critics, when we realize that at virtually every moment we are making decisions—decisions to engage in certain activities and to refrain from engaging in other activities. Consider the fact that you are at this moment reading this book. You certainly do not have to be doing this. Perhaps your professor has made you feel like you must read this, but unless you are being tied down or otherwise physically compelled, it is a choice you are making. You could also be out collecting donations for a certain charity, working in a soup kitchen, or doing any one of a number of things that would clearly produce more overall happiness. According to act-utilitarianism, it seems as if you are acting unethically by reading this book, for you are failing to perform the action that would maximize overall happiness.

This conclusion might itself seem a significant problem for act-utilitarianism; the claim that reading this textbook is morally wrong would strike most people as extremely counterintuitive. But the potential problem is even more serious. If you were to go work at a soup kitchen, for example, that might seem like a very nice, laudable action on your part. (Of course, if that is the alternative that would maximize overall happiness, then it is morally required of you.) After working a full day in the soup kitchen, you would likely be tired and ready to go home. But if going home would produce less overall happiness than staying another hour (and this seems plausible given that you are helping many homeless, often destitute individuals), then you may not leave; it would be required that you stay and continue to help out. An hour later, when you again consider the two alternatives, it seems that the result of the act-utilitarian analysis will be exactly the same; even if you are more fatigued, the overall happiness gained by continuing to help the homeless will likely outweigh the overall happiness of you going home and going to sleep. Even if the calculations eventually do allow you to go home, you will have to go right back first thing in the morning, or else devote your time and efforts to an even more pressing charity (which would produce at least as much happiness). Certainly, this seems very impractical.

The fourth and final criticism to be discussed (there are others in addition to the ones mentioned here) is perhaps the most serious one. Simply stated, the criticism is that the theory at times prescribes acting unethically. One can see why this criticism is potentially very serious indeed; a moral theory is supposed to provide the basis for ethical action, and any moral theory requiring decision-makers to act *unethically* is contradictory to its very purpose and is thus internally incoherent. Examples provide the best way to see why some critics think that act-utilitarianism is flawed in this way.

Imagine you are an office manager in a law firm. Your responsibilities include scheduling the only two administrative assistants, Tom and Sue. (You must make sure they do

not take the same lunch hour or schedule the same week for vacation, for example.) Next Monday is a holiday, and you know from past experience with this holiday that this coming Thursday and Friday will not be busy at all; the lawyers at this firm like to take this opportunity for an extended weekend. Thus, giving each administrative assistant an extra day off, one on Thursday and one on Friday, would cause no problems at all in the firm and would be a nice gesture. But critics of act-utilitarianism argue that in this sort of example a more detailed analysis is needed. Imagine that Tom enjoys days off, but only moderately; he might do laundry or catch up on errands, but the happiness he derives from a day off is limited. Sue, however, derives much more happiness than Tom when she gets a day off, since she likes to go to the beach, the mountains, or into the city. So if you are the office manager, do you necessarily give both Tom and Sue one day off apiece?

The calculations could work out in the following way:

	One Day Off Each	Sue Gets Both Days Off
Sue	+12	+22
Tom	+4	-4
TOTAL	+16	+18

These numerical values are reasonable given the descriptions of the happiness enjoyed by Tom and Sue respectively when given a day off. The first option is straightforward, since Sue derives much more pleasure from a day off. Under the second option, Sue would gain the 12 units of happiness from the Thursday away from work and then another 10 units on Friday. (The difference allows for the principle of diminishing marginal returns, which indicates that less incremental happiness will be derived from each subsequent item of enjoyment. A first piece of pizza, for example, is a bit better than the second, and the second is a bit better than the third, and so on.) Tom, meanwhile, would probably resent not getting one of the two days off, but because a day off is not all that important to him, his resentment would be minimal. Losing 4 units of happiness thus seems appropriate.

Given these values, the morally right thing to do, according to act-utilitarianism, is to give Sue both Thursday and Friday off and tell Tom he has to work. This is what would maximize overall happiness. The criticism, in short, is that act-utilitarianism is flawed because it does not prescribe what is clearly the fair and ethical thing to do, namely scheduling one day off for each administrative assistant.

If this example does not succeed in generating doubts about act-utilitarianism, critics claim to have a better, more convincing one. The following scenario is adapted from one generated by the philosopher Judith Jarvis Thomson. Imagine you are a doctor. At the moment, you are administering a routine physical examination, and the patient is in excellent health. Your mind, however, is really on more troubling matters; there are five patients in the critical care wing of the clinic who are desperately in need of organ transplants. Acceptable donors have not been found, and these particular patients are now critical; they do not have much time. As you are completing the physical exam of the healthy patient, it occurs to you that you do have a certain choice. It would be possible for you to administer a shot to this patient that would cause him to fall asleep and then die painlessly. (You could tell him it is a flu shot or some similar standard precaution.) Because you are very clever, you could concoct an injection that would accomplish this task without anyone being able to discover the true cause of death.

The reason for even considering such an act is that you are an act-utilitarian and you see the possibility of an overall gain in happiness here. You surmise that it might be possible to take various organs out of this healthy patient and redistribute them into the five critical patients as needed—a liver, two kidneys, a heart, and a lung. Because you want to

be sure about the utilitarian calculations, you quickly do some informal research and learn the following: the patient's organs would be excellent matches for the five needy patients (so the likelihood of successful transplants is very high); the patient has no friends or family to speak of, as he just moved here from the coast, "to begin again" as he says (so the unhappiness generated by his death would not be felt by others); each of the five needy patients has children and other family members and friends (so the unhappiness generated by the death of each would be felt by many people). Given all of these circumstances, it seems clear that overall happiness would be maximized by going ahead and killing the innocent, healthy patient in order to save the lives of the five others. After all, there would be a net gain of four lives (five saved, one lost), and consideration of the effects on the happiness of tangential parties (friends and family) adds credence to this conclusion. It seems, then, that according to act-utilitarianism, you as a doctor have a moral obligation to kill your patient. Even if one or two of the transplants were unsuccessful, much more happiness would be produced than would be the case if you refrained from killing the healthy patient.

The example is a bit fanciful, perhaps, but it must be remembered that examples in ethics are often used to make a more general point. The point of the organ transplant example is that act-utilitarianism sometimes prescribes unethical actions, and is thus a very poor (even incoherent) moral theory. The basic structure of the reasoning is as follows. A defensible moral theory must, at the least, prescribe actions that are ethically acceptable. In the case of the organ transplants, act-utilitarianism prescribes an action, killing the patient, that is not ethically acceptable. What follows from these two considerations is the conclusion that act-utilitarianism is not a defensible moral theory.

Utilitarians appeal to two principal sorts of responses to this kind of criticism. First, some continue to defend act-utilitarianism even in the face of examples such as this. It is pointed out that the purpose of a moral theory is to provide a starting point or a foundation for ethical analysis. As discussed in the Introduction, this is thought to be necessary in order to get past some of the main things that cause moral disagreement, such as natural bias, emotional responses, and intuitions (which clearly conflict in many different ethical contexts). In other words, moral theory is important because it forces us to put aside our intuitions; the way we happen to feel about things is descriptive, and should not play a central role in ascertaining the way we should act in different circumstances. With this in mind, utilitarians respond to the fourth objection by pointing out that it *rests* on intuitions, such as our feeling that the doctor who kills her patient acts wrongly. The response, in short, instructs us to get beyond our intuitions and accept the prescriptive implications of a theory (or else to criticize the theory in some other way).

Sometimes, analogies are used to point out the hypocrisy that exists, but is often hidden, in our intuitive reactions to different moral situations. Consider, for example, the decision of the United States to drop the atomic bomb on the city of Hiroshima in August 1945. The strategic aspect of the decision was to bring about an end to the war (and a victory) more quickly. The ethical aspects of the decision included the recognition that many innocent people would be killed, but that many more people would likely be killed if the war were to continue. Estimates of the ratio of the number of people saved by ending the war early (by bombing) to the number of people killed by the bomb varied: some were 3-to-1, some were 5-to-1, some were even higher. The point, though, is that many (though not all) people felt at the time, and in retrospect feel today, that this was a very difficult ethical decision to make. The sense is that the loss of innocent life is always regrettable but is sometimes necessary in order to achieve a greater overall good. Defenders of utilitarianism will point out that this feeling is perfectly in line with their theory and that the decision to drop the bomb—thereby saving, say, five people for every

one person killed—is justified. They will also point out that this is exactly the same reasoning used in the case of the organ transplants—five people will be saved and one will be killed. The reason we happen to feel very negative about that case, and (perhaps) not about the Hiroshima case, is that our intuitions are imperfect. This is exactly the reason they should not be used in analyzing ethical issues, and is the reason moral theory should be used. Even though the analogy is not perfect, act-utilitarians feel that their point is made nonetheless.

In short, the response is that the basis for the fourth criticism is unsound. Other utilitarians, however, are persuaded by this fourth criticism, and are uncomfortable with the idea that we may at times be morally required to perform acts that seem clearly wrong. These latter utilitarians support a second sort of response to the criticism, one that adjusts the way their theory is conceived. They support a version of the theory known as rule-utilitarianism.

### *Rule-Utilitarianism*

Act-utilitarianism calls for the maximization of overall happiness in the circumstances. Rule-utilitarianism differs in that it calls for an analysis that extends beyond the immediate circumstances. Instead, it calls for a closer look at the long-term consequences that would be generated by performing the act in all relevantly similar situations. More formally, according to rule-utilitarianism, the right act conforms to the rule which, when followed, produces at least as much overall happiness over the long run as any alternative rule. More broadly, that rule must then itself be a member of a set of rules that would produce at least as much overall happiness as any alternative set.

The claim made by rule-utilitarians is that their theory will generate rules, and will thus prescribe actions in accordance with those rules, that are much more in line with our intuitions. For example, the rule requiring people to keep their promises is likely justified by rule-utilitarian considerations. A society in which people can rely on others to keep their word will be happier than a society in which promises are broken on a regular basis. Keeping one's word will thus be a fairly stringent moral obligation under rule-utilitarianism; it is a rule that would produce at least as much happiness as any alternative rule pertaining to promise-keeping, and that rule can likely be integrated into a set of rules that would maximize overall societal happiness. Thus, when a lawyer agrees in principle to a settlement in a lawsuit and then goes back on that agreement just before the scheduled trial date (in order to cause confusion on the part of the opposing lawyer), rule-utilitarianism could perhaps be used to show that such an act is unethical.

It will follow that the examples discussed earlier will have more satisfactory conclusions, according to supporters of this version of utilitarianism. In the case of the office manager, act-utilitarianism appeared to recommend giving Sue both days off and making Tom work. Applying rule-utilitarianism, however, would likely generate the conclusion that Sue and Tom should each be given one day off. This is because a general rule stipulating that benefits such as extra days off should be enjoyed by employees *equally* will, in the long run, maximize overall happiness. It will at least produce more happiness than a rule requiring the office manager to engage in various utilitarian calculations every single time a similar situation arises. In the case of the doctor, act-utilitarianism appeared to recommend killing the healthy patient in order to redistribute his organs to the five other patients in need of transplants. Again, though, applying rule-utilitarianism would likely yield a different conclusion. It seems eminently reasonable to think that a rule prohibiting doctors from killing their patients will, in the long run, maximize overall happiness. If pressed, defenders of rule-utilitarianism provide more detailed reasons for thinking

these claims are true. For example, if doctors were sometimes to kill their patients (on those occasions where doing so would be judged to maximize overall happiness in the circumstances), it seems likely that eventually this practice would somehow be discovered. When it is discovered, there will be a widespread tendency not to visit doctors. (Would you go in for a checkup knowing you might be killed?) The result would be a very sickly population and an overall decrease in happiness. Thus, according to this reasoning, following the rule that disallows such killing will produce superior long-run consequences.

It may also be that rule-utilitarianism works well in the context of professional ethics specifically, since rules play such a prominent role in guiding conduct. Indeed, for most ethical issues that arise in the professions, there will often be rules in place that are already applicable. Thus, it is tremendously important to make sure those rules are morally justified. Rule-utilitarianism seems an appropriate moral theory for assessing whether existing rules are indeed justified. Consider the levels of rules discussed in the Introduction. At one level, individual professional organizations typically have their own rules. Lockheed Martin, a well-known international technology corporation, has detailed ethical rules described in its *Code of Conduct*. These guidelines pertain to the safety of employees and customers, accurate record keeping, the process of fairly bidding on contracts, and indeed everything relevant to that sort of professional organization. At a higher level, entire industries often institute ethical rules through broad codes developed by professional societies and organizations. The Society of Professional Journalists, for example, calls for journalists to refuse gifts and favors, avoid stereotyping subjects, and refrain from distorting straightforward facts. If these rules would maximize overall happiness over the long run, then they would be justified by the moral theory of rule-utilitarianism. Even the law itself is a set of rules, and it is not beyond the realm of reason to think that some laws, from time to time, may not be morally justified.

Again, supporters of rule-utilitarianism claim that their theory addresses the fourth criticism of act-utilitarianism, which is that *unethical* actions are prescribed by the theory (such as doctors killing patients). Supporters also maintain that the third criticism, that act-utilitarianism requires too much of people, has been addressed. Instead of constantly working for charity, even at the expense of sleep, it seems logical to think that a rule requiring a balanced approach—allowing for things like sleep, education, and relaxation to be mixed in with the charity work—will maximize happiness over the long run. (Think about the consequences of a society in which we all gave of ourselves to the point of exhaustion.) However, despite the apparent advantages of rule-utilitarianism, it too has been criticized in various ways. For one thing, it does not take care of all the problems. It still requires decision-makers to measure happiness and to predict the future, and critics maintain that it therefore cannot salvage utilitarianism as a workable, defensible moral theory.

In addition, there is a criticism directed toward rule-utilitarianism specifically which maintains that the theory is inherently unstable. The crux of the problem is whether the rules specified by the theory must always be followed or whether exceptions are allowable. Consider the following example. You make an appointment with your professor to discuss the upcoming exam. The appointment is for noon tomorrow at her office, and because you are worried about the exam, you ask her to promise that she will be there. She does promise, and since rule-utilitarianism will support an obligation to keep one's promises, it seems that your professor has an obligation to be there at the appointed time. However, on the way to her meeting with you, she comes across a bad traffic accident. There is only one car involved and only one victim, but he is bleeding to death. No one else is around and there is no way to contact emergency personnel in time to save his life. Your professor knows first aid and can apply a tourniquet to the victim, thereby saving

his life, but in taking the time to do this she will miss her appointment with you. Given these circumstances, her choice is simple: she can break her promise to you in order to stay and save the life of the accident victim, or she can keep her promise to you and leave the victim to bleed to death.

What should she do? Critics allege that regardless of how this question is answered, one cannot be a rule-utilitarian. This is because of the following dilemma. Either the rules are absolute, which means they do not allow for any exceptions, or they are overridable, which means they do allow for exceptions. If they are absolute, then your professor should keep her promise to you; however, since this would allow the victim to die, she would not be maximizing happiness. (It can be assumed that a saved life produces more overall happiness than a discussion of an exam.) Thus, keeping her promise is not a utilitarian solution at all. On the other hand, if the rules are overridable, then your professor should break her promise to you in order to save the man's life; however, while this would be following utilitarianism, it would not be following *rule*-utilitarianism. Allowing for exceptions amounts to saying that the rules should be followed except when the situation is such that breaking the rules would produce more overall happiness—but this is just act-utilitarianism, doing whatever maximizes overall happiness in the circumstances. It is saying that one should keep one's promises as long as doing so would maximize happiness, and then break one's promises when doing so would maximize happiness; this seems to be straightforward act-utilitarianism. Thus, another way of describing this criticism, in addition to saying that it is inherently unstable, is that it collapses back into act-utilitarianism.

### *Wrap-Up of Utilitarianism*

Criticizing utilitarianism is not enough to demonstrate that it should not be used in ethical decision making. In addition to the criticisms, a defense of a "better" theory must be provided. It is much easier to criticize than to defend. (Political campaign advertisements are examples of this; the faults of the other candidate are often highlighted rather than the merits of the candidate himself.) Utilitarians sometimes acknowledge that their theory is not perfect, but they challenge their critics to point to a better one. After all, they say, there is no perfect moral theory; rather, it is the *best* moral theory that is sought. The following theories are attempts to provide an alternative to utilitarianism as a foundation for ethical decision making.

## Further Reading

- Jeremy Bentham, *The Works of Jeremy Bentham*, ed. John Bowring (Edinburgh: William Tate, 1843).
- Alasdair MacIntyre, "Some Consequences of the Failure of the Enlightenment Project," in *After Virtue* (Notre Dame, IN: University of Notre Dame Press, 1984), 62–78.
- John Stuart Mill, *Utilitarianism*, ed. George Sher (Indianapolis, IN: Hackett, 1979).
- G. E. Moore, *Principia Ethica* (London: Cambridge University Press, 1966), chapter one.
- Louis P. Pojman, *Ethics: Discovering Right and Wrong* (Belmont, CA: Wadsworth, 1999), chapter six.
- Samuel Scheffler (ed.), *Consequentialism and Its Critics* (Oxford: Oxford University Press, 1988).

BITS-Pilani



201315  
Central Library

- Michael Slote, *Common-sense Morality and Consequentialism* (London: Routledge and Kegan Paul, 1985), chapter three.
- J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1963).
- Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), chapter five.

## II. Deontological Ethics

In assessing the ethical status of an action or rule, utilitarianism instructs us to focus on the consequences of that action or rule. More specifically, it focuses on the effects on overall happiness. The implication is that other sorts of considerations, such as a person's motives or his overall character, are not relevant. Deontology is an alternative moral theory that differs rather dramatically. First, strict deontologists argue that consequences are completely irrelevant to the ethical status of an action or rule; whereas for utilitarians consequences mean everything, for strict deontologists consequences mean nothing. There are other, more moderate deontologists who do allow consequences to be somewhat relevant, but even they do not agree that there is a good (such as happiness) that must be maximized. It is the rejection of the maximization requirement that deontologists, both strict and moderate, consider to be the main reason that this theory is better than utilitarianism.

The approaches of two leading deontologists will be discussed. The first is a strict deontology described by Immanuel Kant (1724–1804), who is credited with providing the details of the theory. In fact, “deontological ethics” is sometimes thought to be synonymous with “Kantian ethics,” though this equivalence is misguided; one can believe that deontology is the best moral theory without agreeing with the specifics of Kant’s claims. This was true of Sir W. D. Ross (1877–1971), a more moderate deontologist whose views will be discussed after those of Kant.

### *Immanuel Kant*

The purpose of a moral theory, recall, is to provide a way of discovering whether certain actions or policies are ethically right, wrong, or permissible. Consider the following true story. Martell Welch, Jr. attacked Deletha Word because she accidentally dented his car. He stripped and beat her in broad daylight on the Belle Isle bridge in Detroit, Michigan. While he beat her, over forty people watched him and, according to reports, at times cheered for him. No one on the bridge offered any kind of assistance to Word, not even a phone call for help. To escape from Welch, she threw herself over the side of the bridge and drowned. It seems to us, according to our moral vision, that Welch’s actions were ethically wrong, and it also seems that the other people on the bridge were wrong for failing to act on her behalf. A moral theory, though, is supposed to test these sorts of intuitions to see if they reflect moral truth; it is supposed to clarify things, so that we can determine whether our initial moral vision is accurate.

One advantage of utilitarianism is that it is somewhat easier to use than certain other moral theories. The basic principle is straightforward: maximize overall happiness. The way in which Welch’s beating of Word did not maximize overall happiness could be described without undue difficulty. The Kantian version of deontology is more complicated, and understanding why (or even whether) this theory condemns Welch’s actions requires that we proceed carefully.

**Motive** Kant begins his description of his theory with a series of questions. The answer to each question is what Kant calls a “proposition” of morality, and by reviewing these propositions in order we can better understand the theory and how it works.

The first question, which is very basic, is: What makes a person morally good? This seems a reasonable place to begin, and Kant answers this initial question by claiming that the intentions one chooses makes one morally good. When we judge people as morally good or bad, we do not look at whether they happen to achieve their goals. People often fail to accomplish set goals through no fault of their own. You may set yourself the general goal of helping others, and this may mean that when the man next to you has a heart attack, your specific goal is to save his life. Knowing CPR, you can attempt to bring this goal about. It may be, however, that despite your best efforts to save him, the man dies nonetheless. It should be clear that the man’s death does not mean you are morally bad. Instead, it is more plausible to claim that you are morally good because your goal—your intention—was to save his life. This is the idea behind Kant’s first proposition; to assess a person morally, we must look to his intentions.

This first answer gives rise to a second question: What sorts of intentions make one morally good? The answer is Kant’s second proposition of morality. As with the first question, there is a right way and a wrong way to answer this question. The wrong way pertains to happiness. It is a fact of human nature that people are inclined to act so as to make themselves happy. Perhaps you have the intention of doing well in school, and this is because you have the further intention of getting a well-paying and satisfying job, and this is because you have the yet further intention of being happy. This is a perfectly reasonable way of describing a person’s structure of intentions. Kant points out, however, that being happy is not necessarily the same thing as being a morally good person. After all, some things that make you happy may be unethical. Therefore, it is not the intention of bringing about happiness that makes one morally good. Instead, it is acting with the intention of being dutiful—of acting from the motive of duty itself, and not from the (misguided) motive of bringing about happiness. Stated formally, Kant’s second proposition is that a morally good person is motivated to do the right thing just because it is his duty. Thus, acting with a morally good intention (the possession of which makes a person morally good) is the same as acting from the motive of duty.

This second proposition gives rise to a third question: What, exactly, does this mean? What does it mean for a person to intend to act from the motive of duty? Kant’s answer, which is his third proposition, is that acting from the motive of duty is acting out of respect for the moral law. To clarify, “moral law” for Kant refers to that which morality requires. It is analogous to the law handed down by a government; that law is what the government requires of us, and we are to act accordingly. The moral law is what morality itself (objective moral truth) requires of us, and acting out of respect for the moral law means not allowing anything—not happiness, not fear, not love, not even a government’s law—to get in the way of doing what is morally right. This commitment to doing what is right, and being willing to sacrifice happiness along the way, is what Kant means to capture by the notion of “respect” in this third proposition.

To summarize, a person is morally good (performs moral actions) if he acts from a morally good intention, and an intention is morally good if the motive is duty itself, meaning respect for the moral law. This may still be a bit confusing, and certainly it is incomplete, for it leaves out the important issue of how we are to know which actions are morally right in the first place. The process of determining what is morally right, wrong, and permissible is crucially important, of course, and this will be discussed shortly (in the next section). Before proceeding, however, several implications that follow from the discussion of motive should be identified.

First, for Kant, doing the right thing can be a somewhat complex operation. Acting from the right motive is crucial. Some Kantians take this to mean that acting from the right motive is a necessary part of performing the right act—that one cannot perform the right action unless she acts out of respect for the moral law. Others believe that the two notions are separate, that one must perform the morally right action, and in addition, must do so with the right motive (respect for the moral law); on this understanding, it is possible to perform the right action without doing so out of respect for the moral law. Regardless of which interpretation is correct, Kant clearly intended the notion of motive to be the central feature of his moral theory. Second, it follows that appeal to consequences, as indicated previously, is irrelevant. Although certain aspects of the upcoming discussion may appear to incorporate concern for the consequences into Kant's moral reasoning, the appearance is deceptive. Third, it also follows from the emphasis on motive that one's inclinations are to be ignored. Inclinations refer to intuitions, desires, emotions, or any motivations other than respect for the moral law. Kant believes that when determining the ethical course of action, inclinations can get in the way; they can skew one's thinking and lead to misguided conclusions about what is ethical. Inclinations should thus be set aside. A fourth implication, then, is that Kant focuses very narrowly on rationality as the appropriate tool for ethical decision making.

Some of these implications have generated criticisms of Kant's theory, and these will be discussed later. At this point, we are still left with the question of exactly how to determine whether an action or rule is ethically right, wrong, or permissible; the three propositions of morality do not address this. Another way to put the question is that we know we are to act from the motive of duty, and more specifically out of respect for the moral law, but how are we to determine what this moral law is? Kant's answer contains the most famous part of his moral theory, the categorical imperative.

***The Categorical Imperative: Universalizing Maxims*** In ethics, moral theories serve as tools for sharpening our moral vision, for helping us to better see things (ethically) as they really are (objectively). The tool of Kantian deontology focuses on motive, but the more complete account of the theory entails the categorical imperative, which serves two purposes. First, it is a command; it instructs us what to do. Indeed, it tells us what we must do if we are to act ethically, and in this way it tells us what the moral law really is. Second, it is a mechanism for testing actions or rules to see whether they are ethically right, wrong, or permissible. Understanding some of its details will be necessary in order to ascertain what this theory says about Martell Welch's beating of Deletha Word. Kant describes the categorical imperative in several different ways, which he calls formulations. We will be discussing the first two of his formulations, each of which can be helpful when deliberating about ethical issues in the professions.

The first formulation centers on the idea of universalizing maxims. In Kant's words, the command is, "Act only according to that maxim by which you can at the same time will that it should become a universal law." This is straightforwardly a command, and to see how it works as a mechanism for testing actions, several steps can be described. The first step is simply to ask what one's motive for a particular action might be. For example, if we asked Martell Welch why he beat Deletha Word, he might respond by saying, "She made me angry, and beating her made me feel better." His motive in some sense pertains to alleviating the anger she caused him. The second step is to use this motive to ascertain the person's maxim. A maxim is a more general type of motive; it is an individual's reason for acting, but it is expressed as a general rule that applies to all future actions. Welch's motive in the specific case can be expanded into the following general guiding rule: "Whenever anyone angers me, I will beat that person so that I can feel bet-

ter." This is Welch's maxim; it is the general rule that he was following in this particular case. The third step would be to universalize this maxim, which means restating it not just for the individual but for all people (or, as Kant says, for all rational beings). The universalized maxim would then be: "Whenever anyone angers any person, that person who is angry will beat the person who caused the anger."

The fourth step is to assess whether this universalized maxim can be a moral law. If this universalized maxim is consistent—if it can be practiced in the world without any inconsistency—then it can be a moral law. This would mean that the original action performed is not wrong. If, on the other hand, the universalized maxim generates a contradiction and is thus inconsistent, then the original act performed is wrong and may not be performed. In the example, a natural result of the universalized maxim is that there would be a lot of beatings occurring; it is a natural fact that people get angry from time to time, and so numerous beatings would be expected to follow. When people get beaten, they often experience anger (among other emotions), and herein lies the inconsistency. Martell beat Deletha because she made him angry and he wanted to feel better. Thus, on the one hand, he wants his anger to be alleviated. The maxim underlying his action, however, would generate a world in which more anger would come about. This would be true for him as well, since the people *he* angers would beat him, causing him to experience anger at being beaten. In short, he is seeking to lessen his anger by following a general rule (the universalized maxim) that would cause him more anger, and this is straightforwardly inconsistent. Another way to characterize the situation is to say that Martell is acting on a maxim which, when universalized, would be self-defeating, since it would negate the very goal he is trying to accomplish (anger alleviation).

There is certainly some complicated terminology used by Kant, but this first formulation of the categorical imperative can nonetheless be understood fairly clearly through the use of these four basic steps: first, ascertain the individual's motive; second, ascertain the individual's maxim; third, universalize the maxim; and fourth, assess that universalized maxim for consistency. Other examples can be used to make this more clear. For instance, you may have experienced the frustration of being in a traffic jam. Based on this frustration, you may have considered (or even performed) the action of getting onto the shoulder of the highway (or into the empty lane which is closing up ahead) and driving past cars in front of you before merging back into your original lane. Is this action ethically permissible? Answering this question based on the first formulation of the categorical imperative would first require ascertaining your motive, which presumably is to get where you are going faster. The second step is to ascertain your maxim, which is that whenever you want to get where you are going faster, you will "jump ahead" of those in front of you—meaning you will move into the open lane, zoom ahead of others, and cut back into your original lane. As for the third step, the universalized maxim would be that everyone wanting to get where they are going faster will jump ahead in this way. The fourth step reveals that this maxim involves inconsistency. A natural result of everyone trying to jump ahead would be chaos on the road, with everyone trying to constantly change lanes and get ahead of everyone else, and this would mean additional delays. Thus, you would be acting on a maxim which, when universalized, would be self-defeating, since it would negate the very goal you were trying to accomplish (getting where you are going faster).

Kant has several examples of his own demonstrating the operation of this first formulation of the categorical imperative, one of which involves the idea of a lying promise. If a man promises to pay back a loan knowing that he will in fact not pay it back, his action is unethical according to the framework we are using. Without going through all the details, the maxim for making this lying promise would be inconsistent, since a natural result of everyone doing this would be the increased distrust of anyone making a

promise. Thus, in an effort to get another person to trust him, this man would be acting on a maxim which, when universalized, would cause distrust. In the context of the professions, Kantian deontology would deem certain practices unethical for the same reasons. If a business attempts to cheat customers in an effort to make more profit, the universalized maxim would generate the natural result of making clients and customers greatly mistrust businesses, causing them to buy less. This would effectively negate the original goal of increasing profit and thus would be self-defeating.

This discussion of the first formulation of Kant's categorical imperative, which focuses on universalizing maxims, has addressed the ethical status of actions; any action stemming from a maxim that cannot be universalized with consistency is wrong. It should be clear, though, that this same approach can be used to assess the ethical status of rules. A rule would be unethical if it cannot be universalized with consistency and is thus contrary to the moral law. The process would be the same, with the first step assessing the initial motive for adopting the rule rather than the initial motive for performing the action.

Based on the description of how to use this first formulation as a mechanism for testing the ethical status of actions and rules, one might think that an appeal to consequences is, in fact, made by Kant. In the previous examples, we are asked to take into account the natural result of the universalized maxim arrived at in the third step of the analysis. A natural result of the maxim in the Martell Welch example was that there would be numerous beatings, resulting in an increase in anger (contrary to Welch's original goal). A natural result of the maxim in the traffic jam example was that there would be increased traffic problems, resulting in increased delays (contrary to your original goal). These seem, at first glance, to be appeals to consequences, which goes against the initial characterization of Kant as a strict deontologist—as one who believes that consequences are wholly irrelevant to the moral assessment of actions and rules. However, there is no appeal to the *actual* consequences of beating a person who makes you angry or of speeding ahead of other motorists. The analysis called for relies on *idealized* consequences, or perhaps the logical consequences (since the ultimate question is whether there is a contradiction involved). It is not actually the case that every motorist is going to attempt the maneuver described. The question, rather, is what would happen *if* everyone did this, despite the fact that in actuality, that “*if*” clause will not occur. It is in this way that the actual consequences are irrelevant, and it is in this way that Kant is therefore a strict deontologist.

This discussion suggests the ways in which the first formulation of the categorical imperative can be summarized. Perhaps you have heard the adage, “It’s wrong to make yourself an exception to the rule.” This is, essentially, what Kantian deontology is all about—recognizing that a particular rule is appropriate and then following it. More specifically, the motive of duty prescribes that you follow the rule *because* it is the rule. The rules against beating others because you are angry and speeding ahead of other motorists to get where you are going more quickly are seen, through the categorical imperative, to be appropriate. Breaking those rules in order to benefit yourself is thus making yourself an exception to the (justified) rule. Another way to summarize the basic idea is the question, perhaps posed by a parent or teacher when you were growing up, “What if everyone did that?” This question may have been asked in response to your misconduct to get you to think about the process of universalizing, though in a very basic way. Perhaps the best-known summary is the Golden Rule, which instructs us to treat others as we want to be treated. The Golden Rule captures the idea of universalizing fairly well, and it can be a useful way of thinking about the second formulation of the categorical imperative as well.

This approach highlights actions that are unethical. On the other side of the coin, just because a proposed action or rule passes the test set by the first formulation does not

mean it is a moral obligation. One has an obligation to perform an action only if the action is morally right (required); an act or rule passing this test may be just morally permissible. (It should be noted that there is some disagreement among scholars as to whether this particular view about passing the test of the first formulation was indeed Kant's own, though it is generally considered to be consistent with his overall approach.) The practice of opening one's store for business at 8:00 A.M. will pass this first test, but it would be a mistake to conclude that the manager thus has a moral obligation to open her store at that time. One way to hone in on the more precise ethical evaluation is to subject the opposite practice to this test. The follow-up would then be the question of whether failing to open one's store at 8:00 A.M. could be based on a consistent universalized maxim. Surely, the test will reveal that there is also nothing wrong with this (at least with most reasonable motives for opening a store at a given time). By testing both a proposed policy and its opposite, and by learning that nothing is wrong with either, we learn that both are permissible. Thus, a manager may open her store at 8:00 A.M., but doing so is neither required nor forbidden.

Again, the logic of the first formulation focuses on actions or rules that are *wrong*, and when we learn that a certain action X is wrong, it follows logically that *not* doing X is *required*. This is the nature of how we discover moral truth using the first formulation. Thus, learning that it is wrong to beat others when you are angry entails that you are morally required not to beat others when you are angry. What is tricky is that the converse is not necessarily true; when we learn that X is not wrong, it does *not* logically follow that doing X is *required*. Thus, the practice of silently counting to ten when you are angry is not wrong, but this does not entail that you *must* do so. It may turn out that you must do so, but in order to determine this, the opposite (refraining from counting to ten) must be tested to see if it is wrong; if it is wrong, then counting to ten would be required.

This is tricky, not to mention tedious. The second formulation of the categorical imperative may be helpful in enabling us to see, without such mental gymnastics, what is morally required.

***The Categorical Imperative: Persons as Ends*** This second formulation of the categorical imperative is described by Kant in the following way: "Act so that you treat humanity, whether in your own person or in the person of any other, always as an end and never as a means only." It is clear from this that Kant is relying on the importance of the distinction between means and ends, so understanding just what the difference is will be crucial to understanding this second formulation.

In this context, the difference between means and ends is closely related to the difference between people and things. An inanimate object—a "thing"—has value only insofar as someone values it; it does not have value "in itself," meaning it would have no value at all without someone to value it. It is therefore permissible to use things as I see fit in order to accomplish my goals. A computer, for example, is a thing and so has no value in itself. Therefore, since you own your computer, you can use it as a means to your ends—as a tool for helping you achieve what you want. Treating your computer as a mere stepping stone in this way does not morally wrong the computer, for it is not the sort of thing that can be wronged. You can even choose to throw it out the window if you get frustrated with it (as long as no one outside the window will be injured).

In all of these ways, people are different. People do not need to be valued by others in order to have value. Because they are rational beings, people have value in themselves, regardless of how they might be viewed by others. People are therefore not mere things but are "ends in themselves," as Kant says. The ethical relevance of this, according to the

second formulation of the categorical imperative, is that you may not treat another person as a mere means. You may not simply use another person in an effort to further your own goals. People, unlike computers, are not mere stepping stones and may not be treated as such. (Certainly, you may not throw people out the window when you get frustrated with them.) This is Kant's idea behind the moral command that we are always to treat people as ends and never as means only; we must always recognize that people are valuable in themselves (are ends in themselves). He also uses the terminology of *respect* and *dignity* to capture this idea; we are always and everywhere to demonstrate respect for persons, to recognize the inherent dignity they possess (because of their rational capacities). Kant also makes clear that this is why all people are moral equals. Each person is an end in himself to the same degree as all other persons, and so no one is more valuable—more morally important—than anyone else.

Although this is a different formulation, it is still the command of the categorical imperative and is therefore, in a way, redundant. This should become clear upon realizing that the practice of treating people as means only cannot pass the test of the first formulation. Treatment of persons purely in an effort to attain one's own aims—treatment, in other words, that fails to respect persons—does not rest on a maxim that can be universalized. The inconsistency lies in the fact when one treats others as means only, she is operating on the assumption that whatever goal she is trying to accomplish is valuable, which in turn assumes that she views herself as an end. Viewing oneself as an end requires rationality, the ability to reflect on one's goals and act accordingly. Thus, she is basing her view of herself as an end on her rational capacity—but then she is logically required to view all others with rational capacities as ends also (and, hence, not as means only). Failing to do this would be logically inconsistent.

It follows that the examples in the previous sections would come out the same way using this second formulation. Martell Welch's beating of Deletha Word was unethical because he was treating her as a means only, specifically as a means to his end of anger alleviation. By treating her as a thing, he failed to recognize that she is also an end, with goals and values of her own. Your jumping ahead in the traffic jam is unethical because it treats the motorists ahead of you as means only, specifically means to your end of getting to your destination more quickly. By treating them as things, you fail to recognize that they are ends, with goals and values of their own (including, no doubt, the desire to get to their destinations in as timely a manner as possible). In Kant's example, making a lying promise treats the person loaning the money as a means only. This is how, in Kantian deontology, these actions fail to respect others and fail to treat others with the dignity that is owed them.

While technically redundant, the second formulation enables us to further sharpen our moral vision by making it more immediately clear what sorts of actions are required. The decision of whether to open one's store at 8:00 or at 8:30 does not seem to have immediate implications for the treatment of others as ends or as means only. Thus, the conclusion that opening at 8:00 is permissible follows more directly here than it did using the first formulation.

Consider the issue of employee rights in professional ethics. When hiring an employee, a manager's motivation must not simply be to generate more profit for the company. She must recognize that employees are ends in themselves—they are persons, possessing goals and interests of their own—and must not be used as mere stepping stones in the attempt to accomplish some other goal, such as increased profit. When purchasing computers, she *may* treat them as mere means in this way, for computers *are* things, and have value only insofar as people value them. Thus, when the computers get older and become less efficient than the newer models, they may be discarded and the newer ones brought in. Such

a decision has nothing to do with ethics. However, when employees get older and become less efficient than younger individuals who have recently graduated, any decision to “discard” the older ones must be considered carefully. A projected increase in profit is not, in itself, a morally adequate reason for firing long-time employees in order to make room for the younger ones. Indeed, doing so seems to treat those employees as being on a par with computers—as mere means to the end of company profit—and this is what Kant’s second formulation of the categorical imperative would prohibit.

This does not imply that Kantian deontology never allows for the firing of employees. What it does imply is that employees must always be treated as ends. Perhaps reassigning the older employees to other positions or offering early retirement packages would successfully treat them as ends and still accomplish the company’s aims. Of course, these actions may also turn out to be ethically wrong, but this requires a more detailed analysis of how Kantian deontology applies to layoffs. The example should make clear how the theory can be used in the professions.

Other examples can be drawn from the treatment of clients or customers in professional ethics. Consider a baby stroller that is engineered in such a way that its folding mechanism will cause it to collapse onto the child from time to time. Once this problem is discovered, the question will be whether to allow the product to sell or whether to discard the thousands of strollers already produced. After some projections are calculated, the company decides the best business decision is to sell them, since the profits will outweigh the occasional lawsuits generated by injuries to the children. It could be argued, though, that such a strategy would be treating the customers (and their children) as means only, thus violating the categorical imperative. Of course, a strategy other than discarding the thousands of strollers already produced (which would be very costly) may turn out to be acceptable. Perhaps the price could be reduced and the customers informed of the potential risk; they would then be able to make an informed decision about whether to make this purchase, which would save them some money but add an additional risk. If the problem is not discovered until after thousands of strollers have been sold, similar questions will arise about whether to issue a recall (which would be costly), whether to do nothing, or whether some other approach would be ethically permissible. Perhaps distributing free tool kits with instructions for fixing the stroller will be permissible. Because managers must balance the interests not only of customers but also of stockholders and other groups, a decision that treats all parties as ends can be difficult, but this is exactly what is required.

Between the first and second formulations, all professionals should be able to make use of the Kantian approach when establishing or assessing rules, policies, and strategies. Much of this approach can be used without having to engage in the thorny details; internalizing the basic ideas behind the two formulations (namely, the question of what would happen if everyone acted that way and the notion of persons as ends in themselves) will enable one to make many ethically sound decisions as a professional.

**Criticisms of Kant** One possible problem with Kant’s deontology might be referred to as the Ebenezer Scrooge objection. This moral theory instructs us to put aside our inclinations and act on the basis of cold, indifferent reason. It seems, then, that professionals will be most successful in their quest to act ethically if they lose their emotions and become cold and indifferent, much like Scrooge. This is because emotions and inclinations in general can get in the way of acting from the motive of duty and using the categorical imperative, which requires clear, logical reasoning. The problem is that any moral theory that requires us to act like Scrooge, who was initially *unethical*, is flawed from the outset.

Those sympathetic to Kant might respond by saying that the Ebenezer Scrooge objection is a bit unfair. The theory does require acting from the motive of duty, but does not require behaving like Scrooge. The comparison is misleading, since Scrooge (before the visitations of the three spirits) *was* unethical; his motives and his actions were wrong, as demonstrated by his treatment of others. A professional following Kant's theory would not act as Scrooge did. Further, such a professional would not have to adopt a cold or uncaring disposition in carrying out his ethical obligations. While Kantian deontology requires that inclinations be set aside in ethical deliberation, it does not require that they be abandoned altogether. Indeed, Kant says that this would be futile and even harmful; he agrees that we are by nature emotional creatures. The idea, rather, is that we should not be ruled or dominated by our inclinations. This idea, which should sound reasonable, is a far cry from mandating that we all become cold and uncaring.

A second criticism is that setting aside our inclinations, or at least controlling them in the way just described, is not possible. Kant is calling for us to use our rational capacities to control our emotions and inclinations, but if this is not possible then his theory is obviously not very practical. One reason for the criticism is that rationality is a matter of degree, and while we are *somewhat* rational, Kant is mistaken to think we are *so* rational that we can keep our emotions in check as easily as he assumes. Reinforcing this criticism is the observation that lots of us act "irrationally" and act in ways that we know are not very advisable. (Think of various times you have done something you knew was not right, not smart, or not in your own best interests.) The idea, then, is that Kant overestimates the power of rationality and underestimates the power of emotions and inclinations.

The response to this criticism is that it overstates these sorts of difficulties. Put simply, there is no basis for saying that it is impossible to use our rational capacities to control these elements of our personalities. Certainly it is not easy; acting ethically can be very difficult and can cause a great deal of anxiety and internal conflict. But admitting that acting ethically can be difficult is very different from agreeing that it is impossible. Unless someone's ability to reason is not fully developed (such as when one is very young or is mentally handicapped), there exists the ability to recognize what is ethical and to act on it, even if the latter takes a great deal of courage and fortitude.

One of the more frequently cited criticisms is that Kant thought the rules generated by his categorical imperative were absolute, meaning they admitted of no exceptions. Consider that the general moral rule requiring people to keep their promises likely follows from the categorical imperative; breaking promises cannot be universalized (the first formulation) and also treats others, those to whom promises are broken, as means only (the second formulation). Kant would say that promises may therefore *never* be broken, under any circumstances. This creates some troubling results. Recall the example from the discussion of utilitarianism where your professor is faced with the choice of keeping her promise to meet you for an appointment or breaking that promise in order to save the life of an accident victim. The clear intuition is that breaking the promise and saving the victim's life is not only permissible but required, yet Kant's deontological theory, with its absolute moral rules, does not seem to allow this. Kant himself illustrates this point with the example of an innocent man being chased by gangsters who intend to kill him for no good reason. If you know where he is hiding and if the gangsters ask you where he is, you must tell them the truth, even though doing so will result in his wrongful death. This is because the categorical imperative generates a moral rule against lying, and since moral rules are absolute there are no exceptions; one may never lie, not even to save an innocent person's life, and this does not seem right.

A related problem is that if rules are absolute, then there is no way to determine what should be done when the rules conflict. Each of the examples given here can be seen as a

conflict of moral rules. In the first one, the rule requiring promise-keeping is in conflict with the rule requiring assistance of the innocent, and in the second, the rule against lying is in conflict with the rule requiring protection of the innocent. Some mechanism for determining which rule should be followed and which should be overridden must be built into the moral theory itself, yet Kant's theory, with its absolute moral rules, does not seem to have such a mechanism. Thus, one might criticize the theory as being inadequately prescriptive, as well as recommending outcomes that are strongly counterintuitive.

Many consider this to be a very strong objection, and some believe it cannot be responded to adequately. One strategy is to interpret Kant's claims about absoluteness more narrowly. It could be argued that Kant's intention was to suggest that lying for *certain reasons* is always wrong. This would allow for lying in certain situations while prohibiting lying for the wrong reasons, such as to gain personally from the manipulation of others. This sort of response would maintain the absoluteness of Kantian deontology (at least in some sense), though it would have to be conceded that this is a highly dubious interpretation of Kant's actual meaning, especially in light of his example of the gangsters chasing the innocent man. Nonetheless, those sympathetic to the general idea of deontology could still claim that this adjustment improves on Kant, making his theory more plausible.

The immediate problem, of course, is that the question of what makes certain reasons *good* reasons is still up in the air. Moreover, the problem of how to determine which rule should be overridden in conflict situations is not resolved. This first response does not, itself, provide your professor with a way of determining whether saving the victim's life is a good reason for breaking her promise to you. A better response to the problem of absoluteness, at least initially, may be to discard the (apparently) Kantian requirement that moral rules are absolute, and instead adopt a more moderate brand of deontology, one that acknowledges that moral rules requiring such things as honesty and promise-keeping may (and even should) be overridden in certain sorts of situations. Such rules are often referred to as *prima facie*, defined simply as being not absolute. Prima facie rules, then, allow for exceptions. A deontological theory that generates prima facie (rather than absolute) rules would have to take into account at least some aspects of the consequences of those rules, since consideration of consequences seems the only way to determine how to decide between conflicting prima facie moral rules. Hence, it must be some version of moderate (rather than strict) deontology, defined at the beginning of this section as a moral theory emphasizing motive but allowing consequences to still be somewhat relevant. This approach was used by W. D. Ross.

### *Sir W. D. Ross*

Because Ross's theory allows that moral rules can have exceptions, Ross refers to the duties in his theory as *prima facie* duties. Again, prima facie rules are rules that must usually be followed but which have exceptions, and a prima facie duty is a rule, stated in the form of a duty, which usually (but not always) must be carried out. Ross, then, is committed to the idea that there are objective moral rules and that these rules are best stated in terms of the duties we owe each other, but he is not committed to the idea that these duties must always be carried out. There can be extenuating circumstances, such as when prima facie duties conflict. When this occurs one must determine which duty should be overridden. Another way of stating the task is that one must determine which duty is more stringent—which is of more ethical importance—and one question that will arise in the discussion of Ross's theory is whether it provides an adequate basis for assessing how to “weigh” these duties against each other.

John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).

John Rawls, "Justice as Fairness: Political Not Metaphysical," *Philosophy and Public Affairs* 14 (1985).

Michael Sandel, *Liberalism and the Limits of Justice* (Oxford: Blackwell, 1984).

T. M. Scanlon, "Contractualism and Utilitarianism," in Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982).

## IV. Virtue Ethics

### Character

This moral theory differs from the first three in a somewhat significant way. Utilitarianism, deontology, and contractarianism are theories that are designed to help us determine whether rules or actions are morally right, wrong, or permissible. In other words, what all three have in common is that they take rules or actions as the things to be critically assessed. Each has a somewhat different focus—for utilitarianism the focus is on consequences, for deontology it is motivation, and for contractarianism it is hypothetical agreement—but they all yield prescriptions about the moral status of rules or actions. Virtue ethics, on the other hand, focuses not on the moral status of rules or actions but on the moral status of persons, and on individual moral character more specifically. It is a theory whose direct results are about personal moral character. For the purposes of professional ethics, it will therefore reveal what sort of character a professional should have, though we will still want to be able to use it to determine the ethical status of a particular rule or action, and this probably can be done in an indirect way. The additional benefit of this theory, as proclaimed by its adherents, is that it can also tell us, more generally, what sort of person—and thus what sort of professional—we should be.

Most discussions of virtue ethics utilize the work of Aristotle (384–322 B.C.), the famous ancient Greek philosopher whose many writings included the *Nicomachean Ethics*, the *Eudemian Ethics*, and the *Magna Moralia*. Aristotle is credited for assembling the first complete account of virtue ethics, though his teacher Plato (and Plato's teacher, Socrates) certainly had a good deal to say about the topic of virtue. Over the latter half of the twentieth century, a renewed interest in virtue ethics emerged, spurred on by dissatisfaction with the theories of Kant and Mill. One source of dissatisfaction is that those theories emphasize the good of individuals (for example, they generate individual rights and duties) while failing to adequately address the good of the community as a whole. Another criticism has been the treatment of motivation: the theory of utilitarianism suggests that one's motivation for action is irrelevant as long as the right action is performed, and the theory of deontology suggests that one's motivation must be moral duty and nothing else. Contemporary defenders of virtue ethics claim that this theory is more acceptable on these counts.

Although the defenders of virtue ethics have adjusted and fine-tuned Aristotle's basic approach in an effort to make it better, they have not really developed a new version. We will therefore focus on Aristotle's account.

### The Human Function

In focusing on the question of one's moral character, the theory of virtue ethics must first address the question of what it means to be a good person. According to Aristotle,

goodness is to be assessed in terms of function; anything that is judged to be good is judged so because it successfully performs its function. A good knife is one that performs its function well as a knife (which presumably is to cut); a good car is one that performs its function well as a car; a good doctor is one who performs her function well as a doctor. It follows that a good person is one who performs his function well as a person.

So virtue ethics focuses on the notion of good character; having good character means being a good person; and a person is good if he performs his function well as a person. The second question is what the “function” of a person is supposed to be. Aristotle investigated this question by searching for some distinguishing feature of persons, by trying to find an attribute possessed by all (or most) persons and not possessed by other things (knives, cars, plants, animals, and so on). He found the distinguishing feature to be rationality, the ability to use reason and to make important decisions through the use of reason instead of just according to instinct, emotions, and the like. Because of this ability, Aristotle claimed that humans are capable of pursuing the highest form of life, “the good life,” as he called it, or *eudaimonia*. This concept has also been translated as “happiness,” though the meaning is more complex than the simple state-of-mind definition that is usually associated with “happiness”—the definition used in utilitarianism, for example. Happiness for Aristotle was integral to the human function; he claimed that “happiness is an activity of the rational part of the soul in accordance with virtue.”

✓ Putting all this together, we can say that the good person is one who performs his function, which is to strive for the good life by exercising his rationality in accordance with virtue. This statement addresses the two questions noted previously, but it still leaves us wondering about just how we ought to act. It thus gives rise to a third question, which is how, more precisely, we are to act in our day-to-day lives in order to live this “good life.” How, in other words, are we supposed to use our rationality in the way called for by Aristotle? The answer is that we are to act according to the moral virtues. By doing so, we will be acting virtuously and will become virtuous ourselves—that is, we will be fulfilling our function as humans, and will thus acquire good moral character.

### The Virtues

Aristotle’s instruction, then, is to act according to the virtues. In order to do so, we must have a basic understanding of what the virtues are and how they operate. A first point to make is that virtues are ways of acting; they are dispositions, or more accurately habits. Honesty, for example, is a virtue, and the person who acts in accordance with the virtue of honesty is one who has developed the habit of being honest. The generous person—the person who acts in accordance with the virtue of generosity—is one who has gotten to the point where being generous is a habit. Thus, the virtuous person confronted with an ethical situation does not have to engage in theoretical decision-making procedures involving measurements of happiness (utilitarianism), universalized maxims (deontology), or original positions (contractarianism); if she is truly a virtuous person—if she has acquired good moral character—then in most cases she will do the right thing without having to think too much. She will be in the habit of acting in accordance with the moral virtues, and doing the right thing will come naturally to her.

A second point is that Aristotle distinguished the moral virtues, with which we are principally concerned, from the nonmoral virtues. The difference is that the moral virtues are the habits that are necessary for being a morally good person, while the nonmoral virtues are practices that may make us good in other ways. Examples of nonmoral virtues include optimism, cleanliness, and musical talent—nice traits to have, perhaps, but not crucial to being a morally good person. Examples of moral virtues include honesty, gen-

erosity, nonmalevolence, fairness, kindness, and courage (though there is some scholarly debate about which is the appropriate category for courage). It should be clear that the virtues in the latter set are more relevant to moral character; one who has not developed the habits of being honest, kind, fair, and so forth has not yet developed good moral character. (Thus, he is not yet a morally good person—he is not yet fulfilling his function as a person.)

A third point addresses how we are to conceive of each virtue. Again, the moral virtues are habits, but more strongly they are the right kinds of habits, and so the person seeking good moral character must have some way of distinguishing the good habits from the bad ones. Aristotle clarified this through his “doctrine of the mean,” which stipulates that virtues are habits that are in the middle between the two extremes of deficiency and excess. Being virtuous thus entails acting in moderation. For example, being generous means giving, but more precisely it means giving the right amount. One who is stingy does not give enough; stinginess is a deficiency of generosity. One who is extravagant gives too much; extravagance is an excess of generosity. For one who is generous (and recalling the story of Goldilocks and the Three Bears may be helpful here), the amount of giving is “just right.” The moral virtue of courage lies in a mean between the extremes of cowardice (deficiency) and foolhardiness (excess). All other moral virtues are to be understood in the same way. Aristotle’s instruction to strive for the mean is the instruction to avoid the extremes, since acting in extreme ways will not make one a morally good person. It takes practice, but with dedication one can acquire the habit of acting in the mean and avoiding the extremes, and can thus become a morally good person.

According to its proponents, virtue ethics is superior to the other moral theories discussed so far, in part because it allows for realistic motivations. One might even say that in the context of motivation, virtue ethics lies (neatly) in the mean between the extremes of utilitarianism and deontology. Whereas utilitarianism allows that one’s motivation for an action is irrelevant, deontology prescribes that motivation must be moral duty and nothing else. While motivation is relevant according to virtue ethics (thus distinguishing it from utilitarianism), motives can stem from natural desires or emotions (thus distinguishing it from deontology). A person’s natural feeling of compassion can motivate her to contribute to charity, as long as she does not give to the point of extravagance. A person’s natural feeling of pride can motivate her to discuss a personal accomplishment modestly (as long as she does not become boastful); doing so would be practicing the virtue of modesty. In short, the emphasis on moderation allows virtue ethics to be a more “human” moral theory, in that natural feelings and even emotions are allowed to play the right kind of role. Another advantage of virtue ethics is its alleged compatibility with the good of the community as a whole. Here again, one could say that the theory of virtue ethics lies in the mean between two extremes: the overemphasis on individual rights at the expense of the good of the community, and conversely, the overemphasis on the good of the community at the expense of individual rights. It is often said (without going into detail) that by acting virtuously, people will both serve the communal good in the appropriate way and treat individuals as they ought to be treated. The right balance, in other words, will be achieved.

### *Virtue Ethics and Professional Ethics*

Understanding how virtue ethics can be used in the professions may initially seem more difficult than was the case with the previous moral theories, but there are indeed relevant connections. First, virtue ethics centers on character, on being a morally good person, and professionals are, after all, persons. It follows that in order to be an ethical professional,

one must first be an ethical person. A second and related point is that the emphasis on function can be instructive as well. Aristotle was concerned primarily with the human function, but in the professions, we can carry the idea forward and say that a morally good doctor, for instance, is one who performs her function well as a doctor, and that a morally good engineer is one who performs his function well as an engineer. The two points are related in that one must first fulfill one's function as a person before one can understand and fulfill one's function as a certain professional; things cannot work the other way around. As an extreme example to make the point, if one takes himself to be a professional assassin, then in reading Aristotle he might initially be tempted to claim that he will be a morally good assassin if he is successful in performing his function as assassin. The problem is that being an assassin is incompatible with being a morally good person. By first being a good person, he will have the character to understand that it would not be virtuous to become an assassin. Similarly, if one finds oneself in a conflict—between, say, being a good person and being a good lawyer—then one resolves the conflict by looking first at what one should do to be a good person.

\* A third point concerns the idea of moderation. When faced with a difficult ethical decision, a professional may be able to make use of virtue ethics by aiming for the mean and avoiding the extremes. This observation might be relevant to several issues covered in this book, though the details cannot be explored here in any great depth. For example, what is the proper course of action when a company is considering downsizing in order to save money and increase profits? If one takes the attitude that pursuing profit, while important, should be balanced by other ends (such as loyalty to employees), then perhaps downsizing should not occur if profits are already good, even though they could be increased a bit. On the other hand, perhaps downsizing is ethically acceptable if the company is in serious financial trouble. A “balanced” approach that is not too extreme—firing employees for mere marginal gain or retaining employees at the risk of serious financial loss to the company—seems consistent with acting in the mean. Another example, taken from the engineering profession, might pertain to product safety. No product can be perfectly safe. We think products should be “safe enough,” but what, exactly, counts as “safe enough”? In the 1970s, Ford Motor Company was criticized heavily for allowing its Pinto to sell while knowing that impacts of 20 mph could cause the car to explode. Several people died as a result of such explosions. The car “could have been safer,” but this is always the case. Would it have been “safe enough” if explosions were possible only at impacts of at least 30 mph? 40 mph? The safer the car, the more expensive it will be, so where do we draw the line? Aiming for the mean—avoiding the extremes of allowing consumers to buy very unsafe cars on one hand and requiring consumers to spend exorbitant amounts of money for incredibly safe cars on the other hand—might provide the right kind of guidance.

If aiming at the mean makes it possible for someone functioning as a professional to be virtuous, what does the professional hope to gain by being virtuous? Aristotle answers this question, writing that a virtuous person “judges each sort of thing correctly, and in each case what is true appears to him.” In other words, one’s moral vision will become sharper, and this is what we want from a moral theory.

### Criticisms of Virtue Ethics

A first potential problem with virtue ethics concerns its value as an action-guiding framework. A good moral theory, recall, will be adequately prescriptive, meaning it will provide guidance about what ought to be done even in difficult situations. When one finds oneself in a situation in which there is a conflict of virtues, then virtue ethics should pro-

vide some way of adjudicating between them. The criticism, however, is that virtue ethics offers no such means of adjudication; rather than providing a way of determining which virtue ought to be followed, we seem to be left with the very general instruction to just “be virtuous.” This may not help the employee faced with the decision of whether to follow her boss’s instruction to “adjust” certain figures on a financial report in order to make sales look better than they actually are. Virtue ethics may instruct the employee to practice the virtue of loyalty (which would incline her to follow her boss’s instruction) and also to practice the virtue of honesty (which would disincline her to do so), but a mechanism for determining how to “weigh” these against each other is not present.

Proposed solutions to this criticism take several forms, though they amount to the same basic idea. It is sometimes claimed that, in accordance with the doctrine of the mean, the person should aim for moderation and perform the action that is less extreme. Thus, it may be thought that “fudging” a few numbers on a financial report might not be unacceptable if there would be no significant repercussions from doing so and if flatly disobeying one’s boss would cause more extreme problems. Such a scenario may be rare, of course. On the other hand, if the financial report will be used for important purposes (such as setting new sales goals or determining merit-based increases in salary), then adjusting the numbers would be a more extreme action, and the virtuous response would be to refuse the boss’s instruction. This sort of answer appeals to a concept referred to as the “unity of the virtues,” according to which virtues cannot truly conflict—not in a way that produces a genuine moral dilemma. This is because when virtues appear to conflict, there is still a virtuous approach to the situation that can be taken. It is further claimed that if a person is truly virtuous, she will know what to do in these sorts of difficult instances. She will have internalized the virtues to the point where she grasps the unity among them (even if not consciously) and will thus be able to aim toward the mean. Finally, proponents of virtue ethics emphasize that this response demonstrates the superiority of this theory over others, in that virtue ethics does not concern itself with rigid, inflexible rules, including rules for determining what to do when *prima facie* rules conflict. Every situation has its own unique circumstances, they argue, and moral theories that are overly concerned with rules will fail to acknowledge the specific circumstances of each actual situation.

This observation, however, may well be the source of a second criticism. It was noted in the Introduction that rules (such as company and industry practices and codes) play a significant role in professional ethics, and that the use of moral theories in professional ethics will be effective when those theories can help us determine whether rules are ethically right, wrong, or permissible. If virtue ethics focuses on character and not rules, then one might be tempted to conclude that this theory may not be very helpful in professional ethics. Defenders of the theory will respond that it is still possible to use Aristotle’s basic approach to determine whether, for instance, a certain rule is virtuous—whether it lies in the mean and thus helps to promote the good life, or *eudaimonia*. The idea is that even though virtue ethics does not make use of rigid moral rules, it can still be used to ethically assess which actual rules ought to be practiced in the professions. Others might go further and claim that virtues correspond to moral rules, or even are rules themselves. The virtue of honesty would correspond with the moral duty to be honest (the rule requiring honesty), the virtue of fidelity would correspond with the moral duty to keep promises, and so on. Regardless of the precise ways in which this could be worked out, there is no difficulty for the theory of virtue ethics, according to its proponents.

A third criticism is that the theory of virtue ethics commits the so-called “naturalistic fallacy,” which is the illicit move from “what is” (descriptive) to “what ought to be” (prescriptive). The move is illicit because moral theories are supposed to provide ways of

assessing whether “what is” is indeed “what ought to be,” and using the former to determine the latter will prejudice the investigation. In his discussion of the human function, Aristotle notes that humans *are* rational beings (a descriptive claim) and uses this fact to generate ethical claims about how we *ought* (prescriptively) to act. This reasoning, it is alleged, conflates the notions of what we (as humans) happen to be like and what we (as humans) should be like.

A reply to this criticism goes back to the idea of virtue ethics being, for lack of a better term, a more “human” moral theory than others covered in this chapter. Deontology, recall, prescribes moral action by relying on a purely moral motivation, completely separated from our actual motivations, no matter how natural or “human” those actual motivations may be. Contractarianism, recall, prescribes moral action by relying on a scenario, the “original position,” which denies us knowledge of all our contingent attributes and thus denies us everything that makes us who we are as humans. Utilitarianism, recall, relies on an ability to separate and “calculate” total happiness levels and act accordingly, putting aside the personal projects and pursuits that make us human. Those who defend virtue ethics claim that this theory does not require us (unrealistically) to put aside our true selves and proceed according to some detached and unknowable “view from nowhere,” as the philosopher Thomas Nagel puts it. In short, the reply is that the connection made by Aristotle between who we are and who we should be is not undesirable, and that the theory of virtue ethics, which in the search for ethical truth concerning who we should be does not rely on a complete and implausible separation of who we are, is in fact superior to other approaches.

### *Wrap-Up of Virtue Ethics*

Whereas the first three moral theories covered in this chapter maintained that good moral character *followed* from performing the morally appropriate actions, virtue ethics switches the order, the idea being that good moral character *precedes* the performance of right moral action. This difference generates some doubt about the theory’s ability to be adequately action-guiding, but it also may allow some room for “human nature” to play a role in ethics in a way that the previous moral theories do not. While feminist ethics (discussed in the next section) cannot be said to resemble any of the first four to any significant degree, it is probably most like (or more accurately least unlike) virtue ethics because of this allowance for human nature to enter into ethical deliberations.

## Further Reading

- Aristotle, *Nicomachean Ethics* (tr. W. D. Ross), in Richard McKeon (ed.), *The Basic Works of Aristotle* (New York: Random House, 1941).
- John R. Boatright, “Aristotle Meets Wall Street: A Review of Ethics and Excellence,” *Business Ethics Quarterly* 5 (1995): 353–59.
- Daryl Koehn, “A Role for Virtue Ethics in the Analysis of Business,” *Business Ethics Quarterly* 5 (1995): 533–39.
- Thomas Nagel, *The View from Nowhere* (Oxford: Oxford University Press, 1986).
- Louis P. Pojman, *Ethics: Discovering Right and Wrong* (Belmont, CA: Wadsworth, 1999), chapter eight.
- Robert Solomon, *Ethics and Excellence: Cooperation and Integrity in Business* (Oxford: Oxford University Press, 1992).