

HUL 258: Special Topics in Philosophy

Ist Semester, 2024-25

Arudra Burra
Department of Humanities and Social Science
Indian Institute of Technology-Delhi

Course packet II

- 1 “Theorizing about Ethics” (Hugh LaFollette, *Ethics in Practice*)
- 10 “Reading Philosophy” (Hugh LaFollette, *Ethics in Practice*)
- 14 “The Utilitarian Approach” (James Rachels, *Elements of Moral Philosophy*)
- 25 “The Debate over Utilitarianism” (James Rachels, *Elements of Moral Philosophy*)
- 40 “The Good” (Shelley Kagan, *Normative Ethics*)

Theorizing about Ethics

When deciding how to act, we are often faced with uncertainty over, confusions about, or conflicts between, our inclinations, desires, or interests. The uncertainty, confusions, and conflicts can arise even when we are concerned simply to promote our own self-interest. We may not know what is in our best interests: we may have simply adopted some mistaken ideas of our parents, our friends, or our culture. Were our parents Nazis, for example, we may think that maintaining racial purity is our most important personal aim. We may confuse our wants and our interests: we want to manipulate others and therefore infer that close personal relationships are detrimental to our interests. Even when we know some of our interests, we may be unable to rank them by their relative importance: we may assume that wealth is more important than developing character and having close relationships. Other times we may know our interests, but be unsure of how to resolve conflicts between them: I may need to write a paper, yet want to hike the local mountain. Even if I know the best choice, I may not act on it: I may know precisely that it is in my best long-term interest to lose weight, yet inhale that scrumptious pie instead.

These complications show why I can best pursue my interests only if I rationally deliberate about my self interests – if I take the first steps toward theorizing about them. I must sometimes step back and think more abstractly about (a) what it means for something to be an interest (rather than a mere desire), (b) to detect

which objects or behavior or goals are most likely to advance those interests, (c) to understand the interconnections between my interests (e.g., the ways that health enhances my chance of achieving other interests) (d) to find a procedure for coping with conflicts, and (e) to learn how to act on the outcome of rational deliberation. Such theorizing can guide practice: it can help us act more prudently.

Of course, many – perhaps most – actions do not concern simply us; they also concern others, and they do so in myriad ways. Some of my actions may benefit others while other actions may harm them, and I may benefit or harm others directly or indirectly, intentionally or unintentionally. I might directly harm Joe by pushing him. I might push him because I am angry with him or because I want his place in the queue. Or I could indirectly harm Joe, for instance, by landing the promotion he needs to finance nursing care for his dying mother. Or I might offend Joe by privately engaging in what he deems kinky sex. If so, my bedroom antics affect him, although only indirectly, and only because of his particular moral beliefs. Arguably it is inappropriate to say that I harmed Joe in these last two cases, although I did choose to act knowing my actions might affect him (or someone else) in these ways.

In short, in choosing how to act, I should acknowledge that many of my actions affect others, even if only indirectly. In these circumstances, I must choose whether to pursue my self-interest or whether to pursue (or at least

not harm) the interests of others. Other times I must choose to act in ways that may harm some people while benefitting others. I might occasionally find ways to promote everyone's interests without harming anyone's. Occasionally, but not always. Perhaps not even often.

Knowing this does not settle the question of how I should act. It only sets the range within which morality operates. Morality, traditionally understood, involves primarily, and perhaps exclusively, behavior that affects others. I say perhaps because some people (e.g., Kant) think anyone who harms herself, for instance, by squandering her talents or abusing her body, has done something morally wrong. For present purposes, though, we can leave this interesting and important question aside. For what everyone acknowledges is that actions which indisputably affect others fall within the moral arena.

We might disagree about how the fact that an action negatively affects others should shape our decision about how to act. We might also disagree whether and to what extent actions which affect others only indirectly should be evaluated morally. We might further disagree about how to distinguish direct from indirect harm. Nonetheless, if someone's action directly and substantially affects others (either benefits or harms them), then even if we do not yet know whether the action is right or wrong, we can agree that it should be evaluated morally. How we should evaluate it I will discuss in a moment.

But first I should note related but opposite dangers here that we should avoid. One, we might infer from the previous discussion that most moral decisions are complicated or confusing. That is a mistake. For many moral "decisions" are quite easy to make – so easy that we never think about them. No one seriously debates whether morally she should drug a classmate so she can have sex with him, whether she should steal money from her co-workers to finance a vacation on the Riviera, or whether she should knowingly infect someone with AIDS. This is not the stuff of which moral disagreement is made. We know quite well that such actions are wrong. In fact, I dare say that most moral questions are so easily answered that we never ask them. Rather than discuss these obvious "questions," we focus on, think

about, and debate only those that are unclear, those about which there is genuine disagreement.

However, we are also sometimes guilty of assuming that a decision is easy to make, when, in fact, it is not. This opposite extreme is an equally (or arguably more) serious mistake. We may fail to see just the conflicts, confusions, or uncertainties: the issue may be so complicated that we overlook, fail to understand, or do not appreciate how (and how profoundly) our actions affect others. Concern for our self-interest can blind us to the ways our behavior significantly affects others or lead us to give inadequate weight to others' interests. Additionally, our unquestioning acceptance of the moral status quo can make us overlook just how wrong some of our behaviors and institutions are. Widespread acceptability of a practice does not guarantee its correctness.

The Need for Theory

Once we reflect on our thoughts, actions, and choices, we will see that our views are strongly influenced by others. We may think that an action is grossly immoral, but not really know why. Or we may think we know why, only to discover, upon careful examination, that we are just parroting the "reasons" offered by our friends, teachers, parents, or preachers. Of course there is nothing wrong with considering how others think and how they have decided similar moral questions. Indeed, we would be fools not to absorb and benefit from the wisdom of others. Yet anyone who is even faintly aware of history will acknowledge that collective wisdom, like individual wisdom, is sometimes mistaken. Our ancestors held slaves, denied women the right to vote, practiced genocide, and burned witches at the stake. I suspect most of them were generally morally decent people who were firmly convinced that their actions were moral. They acted wrongly because they failed to be sufficiently self-critical. They didn't evaluate their own beliefs; they unquestioningly adopted the outlook of their ancestors, political leaders, teachers, friends, and community. In this they were not unique. This

is a “sin” of which all of us are guilty. The resounding lesson of history is that we must scrutinize our beliefs, our choices, and our actions to ensure that we are informed, consistent, imaginative, unbiased, and not mindlessly repeating the views of others. Otherwise we may perpetrate evils we could avoid, evils for which future generations will rightly condemn us.

One important way to critically evaluate our views is to theorize about ethics: to think about moral issues more abstractly, more coherently, and more consistently. Theorizing is not some enterprise divorced from practice, but is simply the careful, systematic, and thoughtful reflection on practice. Theorizing in this sense will not insulate us from error, but it will empower us to shed ill-conceived, uninformed, and irrelevant considerations. To explain what I mean, let’s think briefly about a matter dear to most students: grades. My grading of students’ work can go awry in at least three different ways.

(1) I might use an inconsistent grading standard. That is, I might use different standards for different students: Joan gets an A because she has a pleasant smile; Ralph, because he works hard; Rachel, because her paper was exceptional. Of course knowing that I should use a unified grading standard does not tell me what standards I should have employed or what grades the specific students should have received. Perhaps they all deserved the A’s they received. However, it is not enough that I accidentally gave them the grades they deserved. I should have given them A’s because they deserved them, not because of these irrelevant considerations. For if I employed irrelevant considerations, I will usually give students the wrong grades even though in these cases, I might have fortuitously given them the appropriate grade.

(2) I might have improper grading standards. It is not enough that I have an invariant standard. After all, I might have a rotten standard to which I adhere unwaveringly. For instance, I might consistently give students I like higher grades than students I dislike. If so, I grade the papers inappropriately, even if consistently.

(3) I might apply the standards inappropriately. I might have appropriate and consistent

grading standards, yet misapply them because I am ignorant, close-minded, exhausted, preoccupied, or inattentive.

I can make parallel “mistakes” in ethical deliberations. For instance:

- 1 I might use inconsistent ethical principles.
- 2 I might have inappropriate moral standards.
- 3 I might apply moral standards inappropriately.

Let us look at each deliberative mistake in more detail:

Consistency We should treat two creatures the same unless they are relevantly different, that is, different in ways that justify treating them differently. Just as students expect teachers to grade consistently, we expect ourselves and others to make moral decisions consistently. The demand for consistency pervades our thinking about ethics. A common strategy for defending our moral views is to claim that they are consistent; a common strategy for criticizing another’s views is to charge that they are inconsistent.

The argumentative role of consistency is evident in the discussion of every practical moral issue. Consider its role in the ABORTION debate. Disputants spend considerable effort arguing that their own positions are consistent while charging that their opponents’ positions are inconsistent. Each side labors to show why abortion is (or is not) relevantly similar to standard cases of murder. Most of those who think abortion is immoral (and likely all of those who think it should be illegal) claim abortion is relevantly similar to murder, while those who think abortion should be legal claim it differs relevantly from murder. What we do not find are people who think abortion is murder and yet wholly moral.

Consistency likewise plays central roles in debates over FREE SPEECH and PATERNALISM AND RISK. Those opposed to censorship often argue that books, pictures, movies, plays, or sculptures that some people want to censor are relevantly similar to other art that most of us do not want censored. They further claim that

Theorizing about Ethics

pornography is a form of speech, and that if it can be prohibited because the majority finds it offensive, then consistency demands that we censor any speech that offends the majority. Conversely, those who claim we can legitimately censor pornography go to some pains to explain why pornography is relevantly different from other forms of speech we want to protect. Both sides want to show that their position is consistent and that their opponent's position is inconsistent.

Although consistency is generally recognized as a requirement of morality, in specific cases it is often difficult to detect if someone is (has been) (in)consistent. Someone may appear to act (in)consistently, but only because we do not appreciate the complexity of her moral reasoning or because we fail to understand the relevant details. As we will see shortly, determining what is and is not morally relevant is often at center stage of many moral debates. Nonetheless, what everyone acknowledges is that *if* someone is being inconsistent, then that is a compelling reason to reject their position unless they can find some way to eliminate that inconsistency.

Correct principles It is not enough to be consistent. We must also employ the appropriate guidelines, principles, standards, or make the appropriate judgments. Theorizing about ethics is one good way to discern the best (most defensible) standards or guidelines, to identify the morally relevant features of our actions, to enhance our ability to make good judgments. Later I discuss how to select and defend these principles – how we determine what is morally relevant.

Correct “application” Even when we “know” what is morally relevant, and even when we reason consistently, we may still make moral mistakes. Consider the ways I might misapply the “rules” prohibiting (a) lying and (b) harming another's feelings. Suppose my wife comes home wearing a gaudy sweater. She wants to know if I like it. Presumably I should neither lie nor intentionally hurt another person's feelings. What, in these circumstances, should I do? There are any number of ways in

which I might act inappropriately. 1) *I may not see viable alternatives*: I may assume, for example, that I must baldly lie or else substantially hurt her feelings. 2) *I may be insufficiently attentive to her needs and interests*: I may over- or under-estimate how much she will be hurt by my honesty (or lack of it). 3) *I may be unduly influenced by self-interest or personal bias*: I may lie not to protect her feelings, but because I don't want her to be angry with me. 4) *I may know precisely what I should do, but be insufficiently motivated to do it*: I may lie because I just don't want the hassle. 5) Or, *I may be motivated to act as I should, but lack the talent or skill to do it*: I may want to be honest, but lack the verbal and personal skills to be honest in a way that will not hurt her feelings.

These are all failings with practical moral significance. We would all be better off if we had the personal traits to avoid these and other moral errors. Ultimately we should learn how to make ourselves more attentive, more informed, and better motivated. However, although these are vitally important practical concerns, they are not the primary focus of most essays in this book. What these authors do here is provide relevant information, careful logical analysis, and a clear account of what they take to be the morally relevant features of practical ethical questions.

Is it just a matter of opinion?

Many of you may find talk of moral standards – and the application of those standards – troubling. You may think – certainly many people talk as if they think – that moral judgments are just “matters of opinion.” All of us have overheard people “conclude” a debate about a contentious moral issue by saying: “Well, it is all just a matter of opinion anyway!” I suspect the real function of this claim is to signal the speaker's desire to terminate the debate – for whatever reason. Perhaps the speaker thinks the other person is irrational and, therefore, that the debate is no longer profitable. Unfortunately this claim seems to imply more. It suggests that since moral judgments are just opinions, then all opinions are equally good (or equally bad). It implies that we cannot criticize

or rationally scrutinize ours (or anyone else's) moral judgments. After all, we don't rationally criticize *mere* opinions.

Is this a defensible implication? I don't see how. Even if no (contentious) moral judgment were indisputably correct, we should not conclude that all moral judgments are equally (un-)reliable. Although we have no clear way of deciding with certainty which actions are best, we have excellent ways of showing that some are defective. We know, for instance, that moral judgments are flawed if they are based on misinformation, shortsightedness, bias, lack of understanding, or wholly bizarre moral principles. Conversely, judgments are more plausible, more defensible, if based on full information, careful calculation, astute perception, and if they have successfully survived the criticism of others in the marketplace of ideas.

Consider the following analogy: no grammatical or stylistic rules will determine precisely the way I should phrase the next sentence. However, from that we should not conclude that I may properly use just any string of words. Some clumps of words are not sentences while some sentences are complete gibberish. Other sentences might be grammatically well formed – even stylish – yet be inappropriate because they are disconnected from the sentences that precede or follow them. All these collections of words are clearly unacceptable in these circumstances, although in some other context(s) the same words might be wholly appropriate. Many other sentences are grammatically well formed, relevant and minimally clear, yet are nonetheless lacking in some way. They might, for instance, be somewhat vague or imprecise. Others may be comprehensible, relevant, and generally precise, yet still be gaudy or at least bereft of style. Some array of others might be wholly adequately, sufficiently adequate so that there is no strong reason to prefer one to the others. Perhaps some would be uniquely brilliant. No grammar book will enable us to make all those distinctions nor could it empower us to clearly identify the best sentence(s). And even if ordinary folks (or even accomplished writers) discussed the merits and demerits of each, we (they) are unlikely to decide that one is uniquely best. Nonetheless, we have no problem distin-

guishing the trashy or the unacceptably vague from the linguistically sublime. In short, we needn't think that one sentence is uniquely good to acknowledge that some are better and some are worse. Likewise for ethics. We may not always know how to act; we may find substantial disagreement about some highly contentious ethical issues. But from that we should not infer that all moral ideas are created equal.

We should also not ignore the obvious fact that circumstances often demand that we act even if there is no (or we cannot discern a) uniquely appropriate moral action. Nonetheless, our uncertainty does not lead us to think that – or act as if – all views were equal. We do not toss a coin to decide whether to remove our parents from life support, whom to marry, which job to take, or whether someone charged with a felony is guilty. We (should) strive to make an informed decision based on the best evidence, and then act accordingly, even though the best evidence will never guarantee certainty. To make an informed decision we should understand the relevant issues, take a longer-term perspective, set aside irrational biases, and inculcate a willingness to subject our tentative conclusions to the criticisms of others.

After all, our actions sometimes do profoundly affect others and circumstances may demand that we act. We should not bemoan our inability to be certain that we have found the uniquely best action; we must simply make the best choice we can. We should, of course, acknowledge our uncertainty, admit our fallibility, and be prepared to consider new ideas, especially when they are supported by strong arguments. However, we have no need to embrace any pernicious forms of relativism. That would be not only misguided, but also a moral mistake.

The Role of Theory

Even when people agree that an issue should be evaluated, at least partly, by criteria of morality, they often disagree about how to evaluate it. Or, using the language of the previous section, people may disagree about the best principles

Theorizing about Ethics

or judgments, about how those are to be interpreted, or about how they should be applied. As a result, two reasonable and decent people can reach wholly different conclusions about the moral appropriateness of the same action. Here is where we clearly require the rational assessment of our actions. We should examine, try to understand, and then evaluate our own and others' reasons in favor of our (their) moral conclusions. After all, people usually have – or at least think they have – reasons for their respective conclusions.

For instance, anti-abortionists argue that abortion is unjustified because the fetus has the same right to life as a normal adult, while pro-abortionists argue that abortion should be legal because the woman has the right to decide what happens in and to her body. Supporters of capital punishment argue that executions deter crime, while opponents argue that it is cruel and inhumane. Those who want to censor pornography claim it degrades women, while supporters argue that it is a form of free speech that should be protected by law.

In giving reasons for their judgments, people usually cite some feature(s) of the action that are thought to explain or undergird that evaluation. This function of reasons is not confined to ethical disagreements. I may justify my claim that “*Fargo* is a good movie” by claiming that it has well-defined characters, an interesting plot, and the appropriate dramatic tension. That is, I identify features of the movie that I think justify my evaluation. The features I cite, however, are not unique to this movie. In giving these reasons I imply that “having well-defined characters” or “having an interesting plot” or “having the appropriate dramatic tension” are important characteristics of good movies, period. That is not to say these are the only or the most important characteristics. Nor is it yet to decide how much weighty these characteristics are. However, it is to say that if a movie has any of these characteristics, then we have a reason to think it is a good movie.

There are three ways you can challenge my evaluation of the movie: you can challenge my criteria, the weight I give those criteria, or my application of the criteria (the claim that

the movie satisfies the criteria). For instance, you could argue that having well-defined characters is not a relevant criterion, that I have given that criterion too much weight, or that *Fargo* does not have well-defined characters. In defense, I could explain why I think it is a relevant criterion, that I have given the criterion the appropriate weight, and that the movie's characters are well developed. At this point we are discussing two related issues arising at “different levels.” We are debating how to evaluate a particular movie, and we are arguing the theoretical merits about competing criteria of good movies.

In like manner, when discussing a practical ethical issue, we are discussing not only that particular issue, we are, whether we realize or not, also engaged in higher level debates about underlying theoretical questions. We do not want to know only whether capital punishment deters crime, we also want to know whether deterrence is morally important, and, if so, just how important. When theorizing reaches a certain level or complexity and sophistication, we can begin to speak of having a theory. Ethical theories are simply formal and more systematic discussions of these second level, theoretical questions. They are philosophers' efforts to identify the relevant moral criteria, the weight or significance of each criterion, and to offer some guidance on how we can determine whether an action satisfies those criteria. In the next section, I will briefly outline some more familiar ethical theories.

But before I do, let me first offer a warning. In thinking about ethical theories, we may be tempted to assume that people who hold the same theory will make the same practical ethical judgements, and that those who make the same practical ethical judgements will embrace the same theory. Not so. That is not true of any evaluative judgements. For instance, two people with similar criteria for good movies may differently evaluate *Fargo*, while two people who loved *Fargo* may have (somewhat) different criteria for good movies. Likewise for ethics. Two people with different ethical theories may nonetheless agree that abortion is morally permitted (or grossly immoral), while two adherents of the same moral theory may differently evaluate

abortion. Knowing someone's theoretical commitments does not tell us precisely what actions she thinks are right or wrong. It tells us only how she thinks about moral issues – about her criteria of relevance and the weight she gives to them.

Main Types of Theory

Two broad classes of ethical theory – consequentialist and deontological – have shaped most people's understanding of ethics. Consequentialists hold that we should choose the available action with the best overall consequences, while deontologists hold that we should act in ways circumscribed by moral rules or rights, and that these rules or rights are defined (at least partly) independently of consequences. Let us look at each in turn. By necessity these descriptions will be oversimplified and somewhat vague. Oversimplified, because I do not have sufficient space to provide a complete account of each. Vague, because even those who embrace these theories disagree about exactly how they should be interpreted. Nonetheless, these descriptions should be sufficient to help you understand the broad outline of each. Then, as you read some of these essays, you will see some ways these theories are developed and wielded in moral debate.

Consequentialism

Consequentialists claim that we are obligated to act in ways that produce the best consequences. It is not difficult to see why this is an appealing theory. First, it relies on the same style of reasoning that we use in making purely prudential decisions. If you are trying to select a college major, you will consider the available options, predict the likely outcomes of each, and determine their relative value. You would then select a major with the best predicted outcome.

Consequentialism uses the same framework, but injects the interests of others into the "equation." When facing a moral decision, I should consider available alternative actions, trace the likely moral consequences of each, and then select the alternative with the best

consequences for all concerned. When stated so vaguely, consequentialism is clearly an appealing moral theory. After all, it seems difficult to deny that achieving the best available outcome would be good. The problem, of course, is deciding which consequences we should consider and how much weight we should give to each. For, until we know that, we cannot know how to reason about morality.

Utilitarianism, the most widely advocated form of consequentialism, has an answer. Utilitarians claim we should choose the option that maximizes "the greatest happiness of the greatest number." They also advocate complete equality: "each to count as one and no more than one." Of course we might disagree about exactly what it means to maximize the greatest happiness of the greatest number; still more we might be unsure about how this is to be achieved. Act utilitarians claim that we determine the rightness of an action if we can decide which action, in those circumstances, would be most likely to promote the greatest happiness of the greatest number. Rule utilitarians, however, reject the idea that moral decisions should be decided case-by-case (see Hooker, *EUTHANASIA*). On their view, we should decide not whether a *particular* action is likely to promote the greatest happiness of the greatest number, but whether a particular *type* of action would, if done by most people, promote the greatest happiness of the greatest number.

Thus, it seems that an act utilitarian might decide that a lie, in a particular case, is justified because it maximizes the happiness of all those concerned, while the rule utilitarian might claim that since everyone's lying would diminish happiness, then it would be best to adopt a strong rule against lying. We should abide by this rule even if, in some particular case, lying might appear to better promote the greatest happiness of the greatest number.

Deontology

Deontological theories are most easily understood in contrast to consequentialist theories. Whereas consequentialists claim we should always strive to promote the best consequences,

Theorizing about Ethics

deontologists claim that our moral obligations – whatever they are – are in some sense and to some degree independent of consequences. Thus, if I have obligations not to kill or steal or lie, those obligations are justified not simply on the ground that following such rules will always produce the best consequences.

That is why many people find deontological theories so attractive. For example, most of us would be offended if someone lied to us, even if the lie produced the greatest happiness for the greatest number. I would certainly be offended if someone killed me, even if my death might produce the greatest happiness for the greatest number (you use my kidneys to save two people's lives, my heart to save someone else's life, etc.). Thus, the rightness or wrongness of lying or killing cannot be explained, the deontologist claims, simply because of its consequences. Of course there is considerable disagreement among deontologists about which rules are true. They also disagree about how to determine these rules. Some deontologists claim abstract reason shows us how we should act (Kant). Others (McNaughton) claim intuitions are our guide. Still others talk about discovering principles that are justified in *reflective equilibrium* (Rawls, e.g., in the selection on ECONOMIC JUSTICE), while some claim we should seek principles that might be adopted by an ideal observer (Arthur in WORLD HUNGER AND INTERNATIONAL JUSTICE).

Alternatives

There are numerous alternatives to these theories. To call them alternatives does not say that they are inferior, only that they have not played as significant a role in the shaping of contemporary ethical thought. Two are especially worth mention since they have become highly influential in the past two decades.

Virtue theory Virtue theory has not been as influential as either deontology or consequentialism in shaping modern ethical thought. Nonetheless, it predates both, at least as a formal theory. It was the dominant theory of the ancient Greeks, reaching its clearest expres-

sion in the work of Aristotle's *Nicomachean Ethics*. For many centuries it was neither discussed nor advocated as a serious competitor. But by the late fifties, it was starting to reappear in the philosophical literature (the history of this re-emergence is traced in the essays reprinted in Crisp and Slote, 1997).

Much of the appeal of virtue theory arises from the perceived failings of the standard alternatives. Deontology and consequentialism, virtue theorists claim, put inadequate (or no) emphasis on the agent – on the ways she should *be*, or the kinds of *character* she should develop. Relatedly, they fail to give appropriate scope to personal judgment and put too much emphasis on following rules (whether deontological or consequentialistic).

Certainly on some readings of deontology and utilitarianism, it sounds as if advocates of these theories believed that a moral decision was the mindless application of a moral rule. The rule says "Be honest," then we should be honest. The rule says: "Always act to promote the greatest happiness of the greatest number," then we need only figure out which action has the most desirable consequences, and then do it. Ethics, then, might be seen to resemble math. The calculations may require patience and care, but they do not depend on judgment.

Many advocates of the standard theories find these objections by virtue theorists telling and over the past two decades, have modified their respective theories to (partially) accommodate them. The result, says Rosalind Hursthouse, is "that the lines of demarcation between these three approaches have become blurred.... Deontology and utilitarianism are no longer perspicuously identified by describing them as emphasizing rules or consequences in contrast to character" (Hursthouse 1999: 4). Both put more emphasis on judgment and on character. For instance, Hill, who is a deontologist, nonetheless describes the proper attitude toward the ENVIRONMENT in a way that emphasizes excellence or character, while Strikwerda and May (SEXUAL AND RACIAL DISCRIMINATION), who do not generally embrace virtue theory, emphasize the need for men to feel shame for their complicity in the rape of women. However, although judgment and character may play in-

creasingly important roles in contemporary versions of deontology or consequentialism, neither play the central role that it does in virtue theory. This is evident, for instance, in Hursthouse's discussion of ABORTION and in the essays in the section on the VIRTUES.

Feminist theory Historically most philosophers were men, men imbued with the sexist outlook of their cultures. Thus, it is not surprising that women's interests, and any perspectives they might have, played no real role in the development of standard ethical theories. The question is: what does that say about those theories? Can we, for instances, merely prune Aristotle's explicit sexism from his theory and still have an Aristotelian theory that is adequate for a less sexist age? Can we prune Kant's sexism and leave a non-sexist but viable deontology?

In the early years of feminism, many thinkers seemed to think so. They claimed that the emphasis in standard ethical theories on justice, equality, and fairness could give women all the argumentative ammunition they needed to claim their rightful place in the public world.

Others were not so sure. For instance, Carol Gilligan (1982) argued that women have different moral experiences and different moral reasoning, and that these differences should be part of any adequate account of morality. She subsequently advocated an "Ethics of Care," which she thought best exemplified women's experience and thinking.

Many subsequent feminists applaud the ways in which the Ethics of Care has critiqued the more standard ethical theories for overlooking, or intentionally eschewing, the experiences and reasoning of women. Yet some of these feminists think that those more traditional theories, especially if augmented by a keen awareness of gender issues and wider concerns with development of people's unique human capacities, can go some distance toward building an adequate ethical theory. Minimally, though, feminist critiques have forced philosophers to reevaluate their theories, and even to rethink exactly what an ethical theory is and what it is supposed to achieve (Jaggar, 2000).

In any event, pay close attention to the ways in which authors discuss issues about RACIAL AND SEXUAL DISCRIMINATION, AFFIRMATIVE ACTION, and ABORTION, as well as several of the essays in the section on FREE SPEECH, and Anderson on Surrogacy (FAMILIES AND REPRODUCTIVE TECHNOLOGY). See whether the reasons used here differ from those employed in other essays. If so, how?

Conclusion

As you read the following essays, you will see how these different ways of thinking about ethics shape our deliberations about particular moral issues. Be alert to these theoretical differences. They will help you better understand the essays. Also pay close attention to the section introductions. These highlight the theoretical issues that play a central role within that section.

Further Reading

- Crisp, R. and Slote, M. A. (eds.) 1997: *Virtue Ethics*. Oxford: Oxford University Press.
- Gilligan, C. 1982: *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Hursthouse, R. 1999: *On Virtue Ethics*. Oxford: Oxford University Press.
- Jaggar, A. M. 2000: Feminist Ethics. In H. LaFollette (ed.), *The Blackwell Guide to Ethical Theory*. Oxford: Blackwell Publishers, pp. 348–74.
- Kant, I. *Grounding of the Metaphysics of Morals*. Indianapolis: Hackett Publishing Co.
- LaFollette, H. 1991: "The Truth in Ethical Relativism." *Journal of Social Philosophy* 20: 146–54.
- LaFollette, H. (ed.) 2000: *The Blackwell Guide to Ethical Theory*. Oxford: Blackwell Publishers.
- McNaughton, D. 1988: *Moral Vision*. Oxford: Blackwell.
- Mill, J. 1861/1979: *Utilitarianism*. Indianapolis: Hackett Publishing Co.
- Rachels, J. 1998: *The Elements of Moral Philosophy* (3rd edn.). New York: McGraw-Hill.
- Scheffler, S. 1992: *Human Morality*. Oxford: Oxford University Press.
- Singer, P. (ed.) 1990: *A Companion to Ethics*. Oxford: Blackwell.

Reading Philosophy

Reading philosophy differs from reading science fiction or the daily newspaper. The subjects are different; the purposes are different; the styles are different. Science fiction attempts to transport us imaginatively to distant worlds of larger-than-life heroes and villains. It aims to entertain us, to divert us from the doldrums of our daily lives, and perhaps even to empower us: having seen the glories or evils of worlds not-yet experienced, we may be better equipped to live in our everyday world. Science fiction achieves these aims by spinning a convincing narrative of creatures living in previously unknown worlds; it evokes our imaginative powers through expressive language.

Newspapers aim to inform us of significant political, social, cultural, economic, and climatic events. Once we are informed we can presumably make better decisions about our leaders, our finances, and our social lives. They typically achieve these aims by giving us the facts, just the facts, and they usually present these facts using a pithy writing style.

Philosophers have neither the simple aims of the journalist nor the airy aims of the science fiction novelist. Their primary function is not to inform or to inspire, but to help us explore competing ideas and the reasons for them. The philosopher achieves these aims by employing a writing style that tends to be neither pithy nor expressive. The style likely

differs from any with which you are accustomed.

Philosophical Language

While the reporter and the novelist write for the public, philosophers usually write for one other. Thus, while most newspapers and some science fiction are written for an eighth grade audience, philosophical essays are written for others with university training. That is why you will need a more robust vocabulary to understand a philosophical essay than you need to understand the latest novel or a column in the local paper. So keep a dictionary handy to look up “ordinary” words you may not yet know. You will also face an additional hurdle with these essays’ vocabularies. Philosophy, like all academic disciplines, employs specialized terms. Some of these are familiar words with specialized meanings; others are words unique to the discipline. To fully grasp philosophical writing, you will need to understand both. Do not despair. Often you can roughly determine the term’s meaning from its context. If, after doing your best, you still cannot understand its meaning, ask your instructor. Most of these words can be explained in a clear, non-technical way. You can also consult the on-line philosophical dictionary (see the link on this book’s supporting web page – www.etsu.edu/Ethics.in.Practice/).

Philosophical writing also tends to be more complex than the writings of reporters and novelists. Occasionally it is more complex than it needs to be: the author may not know how to write clearly. Sometimes the essay *seems* more complex than it is since the author wrote decades or even centuries ago, when most writers penned long, intricate sentences. You can often break down these long sentences into their component parts, e.g., by treating a semicolon as a period. You may also need to reread the essay several times to get a sense of the author's rhythm, much in the way that you may need to listen to a musician several times before you find it easy to appreciate her music.

Often, though, the writing is complex simply because the ideas expressed are complex. We cannot always render profound thoughts into intellectual pabulum. The only way to grasp such essays is to generally improve one's reading skills, in large part by reading and rereading essays until you understand them.

The Centrality of Argument

Philosophical writing is complex also because it contains and evaluates *arguments*. Philosophers forward their own arguments and critique the arguments of others. "Arguments," in this context, have a particular philosophical sense: An argument is a connected series of statements with some central claim the writer is trying to defend (the conclusion), supported by evidence (the premises) the author offers on behalf of the conclusion. The evidence philosophers use varies. They may proffer empirical data, forward imaginative examples, pose suggestions, and critique alternatives. Make certain you have identified the author's conclusion and her premises before you evaluate her work. Do not fall into the trap of judging that an argument is bad simply because you dislike the conclusion.

This tendency to dismiss views we dislike helps explain philosophers' concern with arguments. Each of us is constantly bombarded with claims. Some of these claims are true, some false. Some offer sage wisdom; some, dreadful

advice. How do we distinguish the true from the false, the wise from the stupid – especially when the topic is some controversial moral, political, and social issue? How do we know the proper moral response to abortion, world hunger, homosexuality, and affirmative action? Do we just pick the one we like? The one our parents, preachers, teachers, friends, or society advocate? Often that is exactly what we do. But we shouldn't. Even a cursory glance at history reveals that many horrendous evils were committed by those who embraced their views steadfastly and uncritically. Most Nazis, slave holders, and commanders of Russian Gulags did not think they were immoral; they assumed they were doing the right thing. They simply accepted their society's views without subjecting them to rational scrutiny. That we should not do. At least not if we are responsible individuals. After all, people's lives, welfare, and happiness may depend on our decisions, and the decisions of people like us.

What is our option? We can look for claims supported by the best evidence. We should examine the *reasons* offered for alternative beliefs. Doing so will not insure that we make the best decision, but it will increase the odds that we do. It will lessen the possibility that we make highly objectionable decisions, decisions we will later come to regret. Philosophers offer arguments for their views to help themselves and others make better decisions.

Most people are unaccustomed to scrutinizing arguments. Since most of us were taught to believe what our parents, our priests, our teachers, and our pals told us, we are disinclined to consider the arguments of others seriously, or to rationally criticize our own views. Moreover, although all of us have offered some arguments for our views, we have rarely done so with the care and depth that are the staples of good philosophy. Philosophers strive to offer a clear, unambiguous conclusion supported by reasons that even those disinclined to believe her conclusions are likely to find persuasive. That is not to say that philosophers never make bad arguments or say stupid things. Of course we do. However, it is to say that the explicit aim of philosophy is a *clear, careful assessment of the reasons for and against ours and others' views*.

That is why a key to understanding philosophy is being able to spot arguments, and then to critique them. That is something you will learn, at least in part, by practice.

Looking at Others' Views

Since part of the task of defending one's view is to show that it is rationally superior to alternatives, a philosopher usually not only (a) provides arguments for her view, she will also (b) respond to criticisms of that view, and (c) consider alternative perspectives. Sometimes those other views and criticisms are advocated by a specific philosopher whose work the author cites. Often, though, the view the author discusses is not that of any particular philosopher, but rather the view of some hypothetical advocate of a position (e.g., conservatism or theism or pro-life). This is often double trouble for a student. You may be unfamiliar with the view being discussed. Since you do not know if the view has been accurately represented, you cannot judge if the criticisms (and responses to criticisms) are telling. Worse, you may have trouble distinguishing the author's view from the views of those she discusses. Many students do.

If you read it quickly, and without concentrating, you will probably be confused. However, usually use you can spot this practice if you read the essay carefully. After all, most authors give argumentative road signs that show when she is arguing for a view and when she is stating or discussing someone else's view. Of course the student may also miss these signs because she does not know what to look for. But simply knowing that this is a common argumentative strategy should lessen the difficulty. You can also look for specific cues.

For instance, authors who discuss another's views frequently use the third person to suggest that another person is speaking (or arguing). At other times the author may explicitly say something like "others may disagree..." and then go on to discuss someone else's view. In still other cases the distinction may be more subtle, likely picked up only after carefully reading the essay several times. In the end there is no single

or simple way to distinguish the author's view from other views the author is discussing. However, if you read the essays carefully, using the general strategy just outlined, you will increase the likelihood that you will not be confused.

The Rational Consequences of What We Say

The philosopher's discussion of examples or cases – especially fictional cases – sometimes confuses students. The use of such cases, though, builds upon a central pillar of philosophical argument, namely, that we should consider the implications or rational consequences of our beliefs and actions. The following fictional example explains what I mean. Suppose a teacher gives you an "A" because she likes you, and gives Robert – your worst enemy – an "F" because she dislikes him. You might be ecstatic that you received an "A"; you may also be thrilled to know that your worst enemy failed. However, do you want to say that what the teacher did was morally acceptable? No. There are implications of saying that, implications you would likely be loath to accept.

If you said that the teacher's *reason* for giving those grades was legitimate, you would be saying, in effect, that teachers should be able to give students they like good grades and students they dislike bad grades. Thus, you would be rationally committed to saying that if you had a teacher who disliked you, then she could legitimately fail you. That, of course, is a consequence you are unwilling to accept. Therefore, you (and we) have reason to suspect that your original acceptance of the teacher's grading scheme was inappropriate. This is a common argumentative strategy. Trace the implications – the rational consequences – of a person's *reasons* for action, and then see if you (or others) would be willing to accept those consequences.

A Final Word

These suggestions will not make reading philosophical essays easy. My hope, though, is that it will make it easier. In the end the key to success

is practice. If you have never read philosophical arguments before, you are unlikely to be able to glance at the essay and understand it: you will likely miss the central idea, its relation to alternatives, and you will almost certain fail to comprehend the author's argument. To fully understand the essay, you must read the assignment carefully, and more than once. Most essays are too difficult in style and content for you to grasp in a single reading. Not even most professional philosophers can do that.

Here is a good strategy: Read the essay once. Identify confusing or unusual terms. Try to get a general sense of the argument: what is the point the author wants to establish, what reason does she offer for this claim? What arguments does she discuss? Identify the points about which you are still unclear. After you have a general sense of the essay, reread it more carefully. Strive for a thorough understanding of the argument. Come to class prepared to ask for help clarifying any remaining confusions about the author's views. If you are accustomed to reading an assignment once – and then only quickly – this expectation will seem overly demanding. Yet, it is important that you learn to read carefully and critically.

Herein lies the key to success: persistence and practice. There may be times you find the reading so difficult that you will be tempted to stop, to wait for the instructor to explain it. Yield not to temptation. Press on. It is better and more rewarding to understand the reading for your-

self. Think, for a moment, about what happens when someone "explains" a joke that you could (with time and effort) have understood on your own. It spoils the joke.

Learning to read more complex essays is a skill, and, like any skill, it is not acquired all at once or without effort. Little in life that is valuable is acquired effortlessly. Getting in physical shape requires vigorous exercise and more than a little perspiration. Establishing and maintaining a vibrant relationship requires effort, understanding, and sacrifice. Learning to play a musical instrument does not come quickly, and is, at times, exceedingly frustrating. Learning to read sophisticated essays is no different. If you persist, however, you will find that with time it becomes easier to read and understand philosophical essays. The payoff is substantial and enduring. You will better understand the day's reading assignment, which will most assuredly improve your grade. But more important, you will also expand your vocabulary and hone your reading skills. You will increase your ability to understand more complex and important writing. Most of the world's great books are inaccessible to those with minimal reading and argumentative skills. Learning to read methodically, critically, and in depth will expand your mental horizons. It will increase your understanding of others' views. And it will enhance your ability to refine and defend your own views.

The Utilitarian Approach

Given our present perspective, it is amazing that Christian ethics down through the centuries could have accepted almost unanimously the sententious doctrine that “the end does not justify the means.” We have to ask now, “If the end does not justify the means, what does?” The answer is, obviously, “Nothing!”

JOSEPH FLETCHER, *MORAL RESPONSIBILITY* (1967)

7.1. The Revolution in Ethics

Philosophers like to think their ideas can change the world. Usually, it is a vain hope: They write books that are read by a few other like-minded thinkers, while the rest of humanity goes on unaffected. On occasion, however, a philosophical theory can alter the way people think. Utilitarianism, a theory proposed by David Hume (1711–1776) but given definitive formulation by Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873), is a case in point.

The late 18th and 19th centuries produced an astonishing series of upheavals. The modern nation-state was emerging in the aftermath of the French Revolution and the wreckage of the Napoleonic empire; the revolutions of 1848 showed the continuing power of the new ideas of “liberty, equality, fraternity”; in America, a new country with a new kind of constitution was created, and its bloody civil war was to put an end, finally, to slavery in Western civilization; and all the while the industrial revolution was bringing about a complete restructuring of society.

It is not surprising that in the midst of all this change people might begin to think differently about ethics. The old ways of thinking were very much up in the air, open to challenge. Against this background, Bentham’s argument for a new conception of morality had a powerful influence. Morality, he

urged, is not a matter of pleasing God, nor is it a matter of faithfulness to abstract rules. Morality is just the attempt to bring about as much happiness as possible in this world.

Bentham argued that there is one ultimate moral principle, namely “the Principle of Utility.” This principle requires that whenever we have a choice between alternative actions or social policies, we must choose the one that has the best overall consequences for everyone concerned. Or, as he put it in his book *The Principles of Morals and Legislation*, published in the year of the French Revolution:

By the Principle of Utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question; or what is the same thing in other words, to promote or to oppose that happiness.

Bentham was the leader of a group of philosophical radicals whose aim was to reform the laws and institutions of England along utilitarian lines. One of his followers was James Mill, the distinguished Scottish philosopher, historian, and economist. James Mill’s son, John Stuart Mill, would become the leading advocate of utilitarian moral theory for the next generation, so the Benthamite movement would continue unabated even after its founder’s death.

Bentham was fortunate to have such disciples. John Stuart Mill’s advocacy was, if anything, even more elegant and persuasive than the master’s. In his little book *Utilitarianism* (1861), Mill presents the main idea of the theory in the following way. First, we envision a certain state of affairs that we would like to see come about—a state of affairs in which all people are as happy and well-off as they can be:

According to the Greatest Happiness Principle . . . the ultimate end, with reference to and for the sake of which all other things are desirable (whether we are considering our own good or that of other people), is an existence exempt as free as possible from pain, and as rich as possible in enjoyments.

The primary rule of morality can, then, be stated simply. It is to act so as to bring about this state of affairs, insofar as that is possible:

This, being, according to the utilitarian opinion, the end of human action, is necessarily also the standard of morality, which may accordingly be defined, as the rules and precepts for human conduct, by the observance of which an existence such as has been described might be, to the greatest extent possible, secured to all mankind, and not to them only, but, so far as the nature of things admits, to the whole of sentient creation.

In deciding what to do, we should, therefore, ask what course of conduct would promote the greatest amount of happiness for all those who will be affected. Morality requires that we do what is best from that point of view.

At first glance, this may not seem like such a radical idea; in fact it may seem a mild truism. Who could argue with the proposition that we should oppose suffering and promote happiness? Yet in their own way Bentham and Mill were leading a revolution as radical as either of the other two great intellectual revolutions in the 19th century, those of Marx and Darwin. To understand the radicalness of the Principle of Utility, we have to appreciate what it *leaves out* of its picture of morality: Gone are all references to God or to abstract moral rules “written in the heavens.” Morality is no longer to be understood as faithfulness to some divinely given code or to some set of inflexible rules. The point of morality is seen as the happiness of beings in this world, and nothing more; and we are permitted—even required—to do whatever is necessary to promote that happiness. That, in its time, was a revolutionary idea.

The utilitarians were, as I said, social reformers as well as philosophers. They intended their doctrine to make a difference, not only in thought but in practice. To illustrate this, we will briefly examine the implications of their philosophy for two rather different practical issues: euthanasia and the treatment of nonhuman animals. These matters do not by any means exhaust the practical applications of Utilitarianism; nor are they necessarily the issues that utilitarians would find most pressing. But they do give a good indication of the kind of distinctive approach that Utilitarianism provides.

7.2. First Example: Euthanasia

Matthew Donnelly was a physicist who had worked with X-rays for 30 years. Perhaps as a result of too much exposure, he contracted

cancer and lost part of his jaw, his upper lip, his nose, and his left hand, as well as two fingers from his right hand. He was also left blind. Mr. Donnelly's physicians told him that he had about a year to live, but he decided that he did not want to go on living in such a state. He was in constant pain. One writer said that "at its worst, he could be seen lying in bed with teeth clinched and beads of perspiration standing out on his forehead." Knowing that he was going to die eventually anyway, and wanting to escape this misery, Mr. Donnelly begged his three brothers to kill him. Two refused, but one did not. The youngest brother, 36-year-old Harold Donnelly, carried a .30-caliber pistol into the hospital and shot Matthew to death.

This, unfortunately, is a true story, and the question naturally arises whether Harold Donnelly did wrong. On the one hand, we may assume that he was motivated by noble sentiments; he loved his brother and wanted only to relieve his misery. Moreover, Matthew had asked to die. All this argues for a lenient judgment. Nevertheless, according to the dominant moral tradition in our society, what Harold Donnelly did was unacceptable.

The dominant moral tradition in our society is, of course, the Christian tradition. Christianity holds that human life is a gift from God, so that only he may decide when it will end. The early church prohibited all killing, believing that Jesus' teachings on this subject permitted no exceptions to the rule. Later, some exceptions were made, chiefly to allow capital punishment and killing in war. But other kinds of killing, including suicide and euthanasia, remained forbidden. To summarize the church's doctrine, theologians formulated a rule saying that *the intentional killing of innocent people is always wrong*. This conception has, more than any other single idea, shaped Western attitudes about the morality of killing. That is why we are so reluctant to excuse Harold Donnelly, even though he may have acted from noble motives. He intentionally killed an innocent person; therefore, according to our moral tradition, what he did was wrong.

Utilitarianism takes a very different approach. It would have us ask: Considering the choices available to Harold Donnelly, which one would have the best overall consequences? What action would produce the greatest balance of happiness over unhappiness for all concerned? The person who would be

most affected would, of course, be Matthew Donnelly himself. If Harold does not kill him, he will live on, for perhaps a year, blind, mutilated, and in continuing pain. How much unhappiness would this involve? It is hard to say precisely; but Matthew Donnelly's own testimony was that he was so unhappy in this condition that he preferred death. Killing him would provide an escape from this misery. Therefore, utilitarians have concluded that euthanasia may, in such a case, be morally right.

Although this kind of argument is very different from what one finds in the Christian tradition—as I said before, it depends on no theological conceptions, and it has no use for inflexible “rules”—the classical utilitarians did not think they were advocating an atheistic or antireligious philosophy. Bentham suggests that religion would endorse, not condemn, the utilitarian viewpoint if only its adherents would take seriously their view of God as a *benevolent* creator. He writes:

The dictates of religion would coincide, in all cases, with those of utility, were the Being, who is the object of religion, universally supposed to be as benevolent as he is supposed to be wise and powerful . . . But among the votaries of religion (of which number the multifarious fraternity of Christians is but a small part) there seem to be but few (I will not say how few) who are real believers in his benevolence. They call him benevolent in words, but they do not mean that he is so in reality.

The morality of mercy killing might be a case in point. How, Bentham might ask, could a benevolent God forbid the killing of Matthew Donnelly? If someone were to say that God is kind but that he requires Mr. Donnelly to suffer for the additional year before finally dying, this would be exactly what Bentham means by “calling him benevolent in words, but not meaning that he is so in reality.”

But the majority of religious people do not agree with Bentham, and not only our moral tradition but our legal tradition has evolved under the influence of Christianity. Euthanasia is illegal in every Western nation except Holland. In the United States, it is simply murder, and Harold Donnelly was duly arrested and charged. (I do not know what happened in court, although it is common in such cases for the defendant to be found guilty of a lesser charge and given a light sentence.) What

would Utilitarianism say about this? If, on the utilitarian view, euthanasia is moral, should it also be made legal?

This question is connected with the more general question of what the purpose of the law ought to be. Bentham was trained in the law, and he thought of the Principle of Utility as a guide for legislators as well as for ordinary people making individual moral decisions. The purpose of the law is the same as that of morals: It should promote the general welfare of all citizens. Bentham thought it obvious that if the law is to serve this purpose, it should not restrict the freedom of citizens any more than necessary. In particular, no type of activity should be prohibited unless, in engaging in that activity, one is doing harm to others. Bentham objected to laws regulating the sexual conduct of “consenting adults,” for example, on the grounds that such conduct is not harmful to others, and because such laws diminish rather than increase happiness. But it was Mill who gave this principle its most eloquent expression, when he wrote in his essay *On Liberty* (1859):

The sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, physical or moral, is not a sufficient warrant . . . Over himself, over his own body and mind, the individual is sovereign.

Thus, for the classical utilitarians, laws prohibiting euthanasia are not only contrary to the general welfare, they are also unjustifiable restrictions on people’s right to control their own lives. When Harold Donnelly killed his brother, he was assisting him in concluding his own life in a manner that he had chosen. No harm was caused to anyone else, and so it was none of their business. Most Americans seem to agree with this point of view, at least when it is a practical issue for them. In a 2000 study conducted by the National Institutes of Health, 60% of terminally ill patients said that euthanasia or physician-assisted suicide should be available upon request. Consistent with his philosophy, Bentham himself is said to have requested euthanasia in his final days, although we do not know whether this request was granted.

7.3. Second Example: Nonhuman Animals

The treatment of nonhumans has not traditionally been regarded as presenting much of a moral issue. The Christian tradition says that man alone is made in God's image and that mere animals do not even have souls. Thus the natural order of things permits humans to use animals for any purpose they see fit. St. Thomas Aquinas summed up the traditional view when he wrote:

Hereby is refuted the error of those who said it is sinful for a man to kill dumb animals: for by divine providence they are intended for man's use in the natural order. Hence it is no wrong for man to make use of them, either by killing them or in any other way whatever.

But isn't it wrong to be *cruel* to animals? Aquinas concedes that it is, but he says the reason has to do with human welfare, not the welfare of the animals themselves:

If any passages of Holy Writ seem to forbid us to be cruel to dumb animals, for instance to kill a bird with its young; this is either to remove man's thoughts from being cruel to other men, and lest through being cruel to animals one becomes cruel to human beings; or because injury to an animal leads to the temporal hurt of man, either the doer of the deed, or of another.

Thus people and animals are in separate moral categories. Strictly speaking, animals have no moral standing of their own. We are free to treat them in any way that might seem to our advantage.

When it is spelled out as baldly as this, the traditional doctrine might make one a little nervous: It seems rather extreme in its lack of concern for the animals, many of whom are, after all, intelligent and sensitive creatures. Yet only a little reflection is needed to see how much our conduct is actually guided by this doctrine. We eat animals; we use them as experimental subjects in our laboratories; we use their skins for clothing and their heads as wall ornaments; we make them the objects of our amusement in zoos and rodeos; and, indeed, there is a popular sport that consists in tracking them down and killing them just for the fun of it.

If one is uncomfortable with the theological "justification" for these practices, Western philosophers have offered an

abundance of secular ones. It is said, variously, that animals are not *rational*, that they lack the ability to *speak*, or that they simply are not *human*—and all these are given as reasons why their interests are outside the sphere of moral concern.

The utilitarians, however, would have none of this. On their view, what matters is not whether an individual has a soul, is rational, or any of the rest. All that matters is whether he is capable of experiencing happiness and unhappiness, pleasure and pain. If an individual is capable of suffering, then we have a duty to take that into account when we are deciding what to do, even if the individual in question is not human. In fact, Bentham argues, whether the individual is human or nonhuman is just as irrelevant as whether he is black or white. Bentham writes:

The day *may* come when the rest of the animal creation may acquire those rights which never could have been withholden from them but by the hand of tyranny. The French have already discovered that the blackness of the skin is no reason why a human being should be abandoned without redress to the caprice of a tormentor. It may one day come to be recognized that the number of the legs, the villosity of the skin, or the termination of the *os sacrum* are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or perhaps the faculty of discourse? But a full-grown horse or dog is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day or a week or even a month old. But suppose they were otherwise, what would it avail? The question is not, Can they *reason*? nor Can they *talk*? but, *Can they suffer*?

Because both humans and nonhumans can suffer, we have the same reason for not mistreating both. If a human is tormented, why is it wrong? Because she suffers. Similarly, if a nonhuman is tormented, she also suffers, and so it is equally wrong for the same reason. To Bentham and Mill, this line of reasoning was conclusive. Humans and nonhumans are equally entitled to moral concern.

However, this view may seem as extreme, in the opposite direction, as the traditional view that gives animals no independent moral standing at all. Are animals really to be regarded as

the equals of humans? In some ways Bentham and Mill thought so, but they were careful to point out that this does not mean that animals and humans must always be treated in the same way. There are factual differences between them that often will justify differences in treatment. For example, because humans have intellectual capacities that animals lack, they are able to take pleasure in things that nonhumans cannot enjoy—humans can do mathematics, appreciate literature, and so on. And similarly, their superior capacities might make them capable of frustrations and disappointments that other animals are not able to experience. Thus our duty to promote happiness entails a duty to promote those special enjoyments for them, as well as to prevent any special unhappinesses to which they are vulnerable. At the same time, however, insofar as the welfare of other animals is affected by our conduct, we have a strict moral duty to take that into account, and their suffering counts equally with any similar suffering experienced by a human.

Contemporary utilitarians have sometimes resisted this aspect of the classical doctrine, and that is not surprising. Our “right” to kill, experiment on, and otherwise use animals as we please seems to most of us so obvious that it is hard to believe we really are behaving as badly as Bentham and Mill suggest. Some contemporary utilitarians, however, have produced powerful arguments that Bentham and Mill were right. The philosopher Peter Singer, in a book with the odd-sounding title *Animal Liberation* (1975), has urged, following the principles laid down by Bentham and Mill, that our treatment of nonhuman animals is deeply objectionable.

Singer asks how we can possibly justify experiments such as this one:

At Harvard University R. Solomon, L. Kamin, and L. Wynne tested the effects of electric shock on the behavior of dogs. They placed forty dogs in a device called a “shuttlebox” which consists of a box divided into two compartments, separated by a barrier. Initially the barrier was set at the height of the dog’s back. Hundreds of intense electric shocks were delivered to the dogs’ feet through a grid floor. At first the dogs could escape the shock if they learned to jump the barrier into the other compartment. In an attempt to “discourage” one dog from jumping, the experimenters forced the dog to jump *into* shock 100 times. They said that as the

dog jumped he gave a “sharp anticipatory yip which turned into a yelp when he landed on the electrified grid.” They then blocked the passage between the compartments with a piece of plate glass and tested the same dog again. The dog “jumped forward and smashed his head against the glass.” Initially dogs showed symptoms such as “defecation, urination, yelping and shrieking, trembling, attacking the apparatus” and so on, but after ten or twelve days of trials dogs that were prevented from escaping shock ceased to resist. The experimenters reported themselves “impressed” by this, and concluded that a combination of the plate glass barrier and foot shock were “very effective” in eliminating jumping by dogs.

The utilitarian argument is simple enough. We should judge actions right or wrong depending on whether they cause more happiness or unhappiness. The dogs in this experiment are obviously being caused terrible suffering. Is there any compensating gain in happiness elsewhere that justifies it? Is greater unhappiness being prevented, for other animals or for humans? If not, the experiment is not morally acceptable.

We may note that this style of argument does not imply that all such experiments are immoral—it suggests judging each one individually, on its own merits. The experiment with the dogs, for example, was part of a study of “learned helplessness,” a topic that psychologists regard as very important. Psychologists say that finding out about learned helplessness will lead to long-term benefits for the mentally ill. The utilitarian principle does not, by itself, tell us what the truth is about particular experiments; but it does insist that the harm done to the animals *requires justification*. We cannot simply assume, because they are not human, that anything goes.

But criticizing such experiments is too easy for most of us. Because we do not do such research, we may feel superior or self-righteous. Singer points out, however, that none of us is free of blame in this area. We are all involved in cruelty just as serious as that perpetrated in any laboratory, because we all (or, at least most of us) eat meat. The facts about meat production are at least as harrowing as the facts about animal experimentation.

Most people believe, in a vague way, that while the slaughterhouse may be an unpleasant place, animals raised for food are otherwise treated well enough. But, Singer points out, noth-

ing could be further from the truth. Veal calves, for example, spend their lives in pens too small to allow them to turn around or even to lie down comfortably—but from the producers' point of view, that is good, because exercise toughens the muscles, which reduces the "quality" of the meat; and besides, allowing the animals adequate living space would be prohibitively expensive. In these pens the calves cannot perform such basic actions as grooming themselves, which they naturally desire to do, because there is not room for them to twist their heads around. It is clear that the calves miss their mothers, and like human infants they want something to suck: They can be seen trying vainly to suck the sides of their stalls. In order to keep their meat pale and tasty, they are fed a liquid diet deficient in both iron and roughage. Naturally they develop cravings for these things. The calf's craving for iron becomes so strong that if allowed to turn around, it will lick at its own urine, although calves normally find this repugnant. The tiny stall, which prevents the animal from turning, solves this "problem." The craving for roughage is especially strong, since without it the animal cannot form a cud to chew. It cannot be given any straw for bedding, since the animal would be driven to eat it, and that would affect the meat. So for these animals, the slaughterhouse is not an unpleasant end to an otherwise contented existence. As terrifying as the process of slaughter is, for them it may actually be a merciful release.

Once again, given these facts, the utilitarian argument is simple enough. The system of meat production causes great suffering for the animals. Because we do not need to eat them—vegetarian meals are also tasty and nourishing—the good that is done does not, on balance, outweigh the evil. Therefore, it is wrong. Singer concludes that we should become vegetarians.

What is most revolutionary in all this is simply the idea that the interests of nonhuman animals *count*. We normally assume, as the dominant tradition of our society teaches, that human beings alone are worthy of moral consideration. Utilitarianism challenges that basic assumption and insists that the moral community must be expanded to include all creatures whose interests are affected by what we do. Human beings are in many ways special; and an adequate morality must acknowledge that. But it is also true that we are only one species among many inhabiting this planet; and morality must acknowledge that as well.

The Debate over Utilitarianism

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

JOHN STUART MILL, *UTILITARIANISM* (1861)

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

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

8.1. The Classical Version of the Theory

Classical Utilitarianism, the theory of Bentham and Mill, can be summarized in three propositions: First, actions are to be judged right or wrong solely by virtue of their consequences. Nothing else matters. Second, in assessing consequences, the only thing that matters is the amount of happiness or unhappiness that is created. Everything else is irrelevant. Third, each person's happiness counts the same. As Mill put it,

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

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

The appeal of this theory to philosophers, economists, and others who theorize about human decision making has been enormous. The theory continues to be widely accepted, even though it has been challenged by a number of apparently devastating arguments. These antiutilitarian arguments are so nu-

merous, and so persuasive, that many have concluded the theory must be abandoned. But the remarkable thing is that so many have not abandoned it. Despite the arguments, a great many thinkers refuse to let the theory go. According to these contemporary utilitarians, the antiutilitarian arguments only show that the classical theory needs to be improved; they say the basic idea is sound and should be preserved, but recast into a more satisfactory form.

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

8.2. Is Happiness the Only Thing That Matters?

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

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

A promising young pianist’s hands are injured in an automobile accident so that she can no longer play. Why is this bad for her? Hedonism would say it is bad because it causes her unhappiness. She will feel frustrated and upset whenever she thinks of what might have been, and *that* is her misfortune. But this way of explaining the misfortune seems to get things the wrong way around. It is not as though, by feeling unhappy, she has made

an otherwise neutral situation into a bad one. On the contrary, her unhappiness is a rational response to a situation that *is* unfortunate. She could have had a career as a concert pianist, and now she cannot. That is the tragedy. We could not eliminate the tragedy just by getting her to cheer up.

You think someone is your friend, but he ridicules you behind your back. No one tells you, so you never know. Is this unfortunate for you? Hedonism would have to say no, because you are never caused any unhappiness. Yet we feel there is something bad going on. You think he is your friend, and you are “being made a fool,” even though you are unaware of it and you suffer no unhappiness.

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

In this way, Hedonism misunderstands the nature of happiness. Happiness is not something that is recognized as good and sought for its own sake, with other things desired only as a means of bringing it about. Instead, happiness is a response we have to the attainment of things that we recognize *as* good, independently and in their own right. We think that friendship is a good thing, and so having friends makes us happy. That is very different from first setting out after happiness, then deciding that having friends might make us happy, and then seeking friends as a means to this end.

For this reason, there are not many hedonists among contemporary philosophers. Those sympathetic to Utilitarianism have therefore sought a way to formulate their view without assuming a hedonistic account of good and evil. Some, such as the English philosopher G. E. Moore (1873–1958), have tried to compile short lists of things to be regarded as good in themselves. Moore suggested that there are three obvious intrinsic goods—pleasure, friendship, and aesthetic enjoyment—and that right actions are those that increase the world’s supply of these things. Other utilitarians have bypassed the question of how many things are good in themselves, leaving it an open

question and saying only that right actions are the ones that have the best results, however that is measured. Still others bypass the question in another way, holding only that we should act so as to maximize the satisfaction of people's *preferences*. It is beyond the scope of this book to discuss the merits or demerits of these varieties of Utilitarianism. I mention them only in order to note that, although the hedonistic assumption of the classical utilitarians has largely been rejected, contemporary utilitarians have not found it difficult to carry on. They do so by urging that Hedonism was never a necessary part of the theory in the first place.

8.3. Are Consequences All That Matter?

The idea that only consequences matter is, however, a necessary part of Utilitarianism. The theory's most fundamental idea is that in order to determine whether an action would be right, we should look at *what will happen as a result of doing it*. If it were to turn out that some other matter is also important in determining rightness, then Utilitarianism would be undermined at its very foundation.

Some of the most serious antiutilitarian arguments attack the theory at just this point: They urge that various other considerations, in addition to utility, are important in determining right and wrong. Here are three such arguments.

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

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

This is a fictitious example, of course, although it was obviously inspired by the lynch-law that prevailed at one time in some

parts of the United States. In any case, the argument is that if someone were in this position, then on utilitarian grounds he should bear false witness against the innocent person. This might have some bad consequences—the innocent man might be executed—but there would be enough good consequences to outweigh them: The riots and lynchings would be stopped. The best outcome would be achieved by lying; therefore, according to Utilitarianism, lying is the thing to do. But, the argument continues, it would be wrong to bring about the execution of an innocent person. Therefore, Utilitarianism, which implies it would be right, must be incorrect.

According to the critics of Utilitarianism, this argument illustrates one of the theory's most serious shortcomings: namely, that it is incompatible with the ideal of justice. Justice requires that we treat people fairly, according to their individual needs and merits. McCloskey's example illustrates how the demands of justice and the demands of utility can come into conflict. Thus, an ethical theory that says utility is the whole story cannot be right.

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

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

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

Later that month, Story advised appellant that the pictures did not come out and that he had destroyed them. Instead, Story circulated these photographs among the personnel of the Chino police department. In April, 1960,

two other officers of that police department, appellee Louis Moreno and defendant Henry Grote, acting under color of their authority as such, and using police photographic equipment located at the police station made additional prints of the photographs taken by Story. Moreno and Grote then circulated these prints among the personnel of the Chino police department.

Ms. York brought suit against these officers and won. Her legal rights had clearly been violated. But what of the *morality* of the officers' behavior? Utilitarianism says that actions are defensible if they produce a favorable balance of happiness over unhappiness. This suggests that we consider the amount of unhappiness caused to Ms. York and compare it with the amount of pleasure taken in the photographs by Officer Story and his cohorts. It is at least possible that more happiness than unhappiness was caused. In that case, the utilitarian conclusion apparently would be that their actions were morally all right. But this seems to be a perverse way of thinking. Why should the pleasure afforded Story and his cohorts matter at all? Why should it even count? They had no right to treat Ms. York in this way, and the fact that they enjoyed doing so hardly seems a relevant defense.

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

The moral to be drawn from this argument is that Utilitarianism is at odds with the idea that people have *rights* that may not be trampled on merely because one anticipates good results. In these cases, it is Ms. York's right to privacy that is violated; but it would not be difficult to think of similar cases in which other rights are at issue—the right to freedom of religion, to free speech, or even the right to life itself. It may

happen that good purposes are served, from time to time, by violating these rights. But we do not think that our rights should be set aside so easily. The notion of a personal right is not a utilitarian notion. Quite the opposite: It is a notion that places limits on how an individual may be treated, regardless of the good purposes that might be accomplished.

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

There is an important general lesson to be learned from this argument. Why is Utilitarianism vulnerable to this sort of criticism? It is because the only kinds of considerations that the theory holds relevant to determining the rightness of actions are considerations having to do with the future. Because of its exclusive concern with consequences, Utilitarianism has us confine our attention to what *will happen* as a result of our actions. However, we normally think that considerations about the past are also important. (The fact that you promised your friend to meet her is a fact about the past.) Therefore, Utilitarianism seems to be faulty because it excludes backward-looking considerations.

Once we understand this point, other examples of backward-looking considerations come easily to mind. The fact that someone did not commit a crime is a good reason why he should not be punished. The fact that someone once did you a favor may

be a good reason why you should now do him a favor. The fact that you did something to hurt someone may be a reason why you should now make it up to her. These are all facts about the past that are relevant to determining our obligations. But Utilitarianism makes the past irrelevant, and so it seems deficient for just that reason.

8.4. Should We Be Equally Concerned for Everyone?

The final component of utilitarian morality is the idea that we must treat each person's welfare as equally important—as Mill put it, that we must be “as strictly impartial as a disinterested and benevolent spectator.” This sounds plausible when it is stated abstractly, but it has troublesome implications. One problem is that the requirement of “equal concern” places too great a demand on us; another problem is that it disrupts our personal relationships.

The Charge That Utilitarianism Is Too Demanding. Suppose you are on your way to the theater when someone points out that the money you are about to spend could be used to provide food for starving people or inoculations for third-world children. Surely, those people need food and medicine more than you need to see a play. So you forgo your entertainment and give the money to a charitable agency. But that is not the end of it. By the same reasoning, you cannot buy new clothes, a car, a computer, or a camera. Probably you should move into a cheaper apartment. After all, what is more important—your having these luxuries or children having food?

In fact, faithful adherence to the utilitarian standard would require you to give away your resources until you have lowered your own standard of living to the level of the neediest people you could help. We might admire people who do this, but we do not regard them as simply doing their duty. Instead, we regard them as saintly people whose generosity goes *beyond* what duty requires. We distinguish actions that are morally required from actions that are praiseworthy but not strictly required. (Philosophers call the latter *supererogatory* actions.) Utilitarianism seems to eliminate this distinction.

But the problem is not merely that Utilitarianism would require us to give up most of our material resources. Equally important, abiding by Utilitarianism's mandates would make it impossible for us to carry on our individual lives. Each of our lives includes projects and activities that give it character and meaning; these are what make our lives worth living. But an ethic that requires the subordination of everything to the impartial promotion of the general welfare would require us to abandon those projects and activities. Suppose you are a cabinet-maker, not getting rich but making a comfortable living; you have two children that you love; and on weekends you like to perform with an amateur theater group. In addition you are interested in history and you read a lot. How could there be anything wrong with this? But judged by the utilitarian standard, you are leading a morally unacceptable life. After all, you could be doing a lot more good if you spent your time in other ways.

Personal Relationships. In practice, none of us is willing to treat all people as equals, for it would require that we abandon our special relationships with friends and family. We are all deeply partial where our friends and family are concerned. We love them and we go to great lengths to help them. To us, they are not just members of the great crowd of humanity—they are special. But all this is inconsistent with impartiality. When you are impartial, intimacy, love, affection, and friendship fly out the window.

The fact that Utilitarianism undermines our personal relationships seems to many critics to be its single greatest fault. Indeed, at this point Utilitarianism seems to have lost all touch with reality. What would it be like to be no more concerned for one's husband or wife than for strangers whom one has never met? The very idea is absurd; not only is it profoundly contrary to normal human emotions, but the institution of marriage could not even exist apart from understandings about special responsibilities and obligations. Again, what would it be like to treat one's children with no greater love than one has for strangers? As John Cottingham puts it, "A parent who leaves his child to burn, on the ground that the building contains someone else whose future contribution to the general welfare promises to be greater, is not a hero; he is (rightly) an object of moral contempt, a moral leper."

8.5. The Defense of Utilitarianism

These arguments add up to an overwhelming indictment of Utilitarianism. The theory, which at first seemed so progressive and commonsensical, now seems indefensible: It is at odds with such fundamental moral notions as justice and individual rights, and it seems unable to account for the place of backward-looking reasons in justifying conduct. It would have us abandon our ordinary lives and spoil the personal relationships that mean everything to us. Not surprisingly, the combined weight of these arguments has prompted many philosophers to abandon the theory altogether.

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

The First Line of Defense: Fanciful Examples Don't Matter.

The first line of defense is to argue that the antiutilitarian arguments make unrealistic assumptions about how the world works. The arguments about rights, justice, and backward-looking reasons share a common strategy. A case is described, and then it is said that from a utilitarian point of view a certain action is required—bearing false witness, violating someone's rights, or breaking a promise. It is then said that these things are not right. Therefore, it is concluded, the utilitarian conception of rightness cannot be correct.

But this strategy succeeds only if we agree that the actions described really would have the best consequences. But why should we agree with that? In the real world, bearing false witness does *not* have good consequences. Suppose, in the case described by McCloskey, the “utilitarian” tried to incriminate the innocent man in order to stop the riots. He probably would not succeed; his lie might be found out, and then the situation would be even worse than before. Even if the lie did succeed, the real culprit would remain at large, to commit additional crimes. Moreover, if the guilty party were caught later on, which is always a possibility, the liar would be in deep trouble, and confidence in the criminal justice system would be undermined. The moral is that although one might *think* that one can bring about the best consequences by such behavior, one can by no means be certain of it. In fact, experience teaches the contrary: Utility is not served by framing innocent people.

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

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

So that is the first line of defense. How effective is it? Unfortunately, it contains more bluster than substance. While it can plausibly be maintained that *most* acts of false witness and the like have bad consequences in the real world, it cannot reasonably be asserted that *all* such acts have bad consequences. Surely, at least once in a while, one can bring about a good result by doing something that moral common sense condemns. Therefore, in at least some real-life cases Utilitarianism will come into conflict with common sense. Moreover, even if the antiutilitarian arguments had to rely exclusively on fictitious examples, those arguments would nevertheless retain their power; for showing that Utilitarianism has unacceptable consequences in hypothetical cases is a valid way of pointing up its theoretical defects. The first line of defense, then, is weak.

The Second Line of Defense: The Principle of Utility Is a Guide for Choosing Rules, Not Individual Acts. The second line of defense admits that the classical version of Utilitarianism is inconsistent with moral common sense and proposes to save the theory by giving it a new formulation which *will* be in line with

our commonsense evaluations. In revising a theory, the trick is to identify precisely which of its features is causing the trouble and to change that, leaving the rest of the theory undisturbed. What is it about the classical version that generates all the unwelcome results?

The troublesome aspect of classical Utilitarianism, it was said, is its assumption that *each individual action* is to be evaluated by reference to the Principle of Utility. If on a certain occasion you are tempted to bear false witness, the classical version of the theory says that whether it would be wrong is determined by the consequences of *that particular lie*; similarly, whether you should keep a promise depends on the consequences of *that particular promise*; and so on for each of the examples we have considered. This is the assumption that caused all the trouble; it is what leads to the conclusion that you can do any sort of questionable thing if it has the best consequences.

Therefore, the new version of Utilitarianism modifies the theory so that individual actions will no longer be judged by the Principle of Utility. Instead, we first ask what *set of rules* is optimal, from a utilitarian point of view. What rules should we prefer to have current in our society, if the people in our society are to flourish? Individual acts are then judged right or wrong according to whether they are acceptable or unacceptable by those rules. This new version of the theory is called *Rule-Utilitarianism*, to distinguish it from the original theory, now commonly called *Act-Utilitarianism*. Richard Brandt was perhaps the most prominent defender of Rule-Utilitarianism; he suggested that “morally wrong” means that an action

would be prohibited by any moral code which all fully rational persons would tend to support, in preference to all others or to none at all, for the society of the agent, if they expected to spend a lifetime in that society.

Rule-Utilitarianism has no difficulty coping with the anti-utilitarian arguments. An act-utilitarian, faced with the situation described by McCloskey, would be tempted to bear false witness against the innocent man because the consequences of *that particular act* would be good. But the rule-utilitarian would not reason in that way. He would first ask, “What general rules of conduct tend to promote the greatest happiness?” Suppose we imagine two societies, one in which the rule “Don’t bear false

witness against the innocent” is faithfully adhered to, and one in which this rule is not followed. In which society are people likely to be better off? From the point of view of utility, the first society is preferable. Therefore, the rule against incriminating the innocent should be accepted, and by appealing to this rule, we conclude that the person in McCloskey’s example should not testify against the innocent man.

Analogous reasoning can be used to establish rules against violating people’s rights, breaking promises, lying, and all the rest. Rules governing personal relationships—requiring loyalty to friends, loving care of one’s children, and so on—can also be established in this manner. We should accept such rules because following them, as a regular practice, promotes the general welfare. But once having appealed to the Principle of Utility to establish the rules, we do not have to invoke the principle again to determine the rightness of particular actions. Individual actions are justified simply by appeal to the already-established rules.

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

The Third Line of Defense: “Common Sense” Can’t Be Trusted. Finally, a small group of contemporary utilitarians has had a very different response to the antiutilitarian arguments. Those arguments point out that the classical theory is at odds with ordinary notions of justice, individual rights, and so on; and this group responds: “So what?” In 1961 the Australian philosopher J. J. C. Smart published a monograph entitled *An Outline of a System of Utilitarian Ethics*; reflecting on his position in that book, Smart said:

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

Our moral common sense is, after all, not necessarily reliable. It may incorporate various irrational elements, including preju-

dices absorbed from our parents, our religion, and the general culture. Why should we simply assume that our feelings are always correct? And why should we reject a plausible, rational theory of ethics simply because it conflicts with those feelings? Perhaps it is the feelings, not the theory, that should be discarded.

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

And consider again the objection that Utilitarianism is too demanding because it would require us to use our resources to feed starving children rather than go to movies and buy cars and cameras. Is it unreasonable to believe that continuing our affluent lives is less important than those children?

On this way of thinking, Act-Utilitarianism is a perfectly defensible doctrine and does not need to be modified. Rule-Utilitarianism, by contrast, is an unnecessarily watered-down version of the theory, which gives rules a greater importance than they merit. There is a serious problem with Rule-Utilitarianism, which can be brought out if we ask whether its rules have *exceptions*. After the rule-utilitarian's "ideal social code" has been established, are its rules to be followed no matter what? There will inevitably be cases where an act that is prohibited by the code would nevertheless maximize utility, maybe even by a considerable amount. Then what is to be done? If the rule-utilitarian says that in such cases we may violate the code, it looks like he has fallen back into act-utilitarianism. On the other hand, if he says that we may not do the "forbidden" act, then, as Smart puts it, the utilitarian's original concern for promoting welfare has been replaced by an irrational "rule worship." What sort of utilitarian would allow the sky to fall for the sake of a rule?

Act-Utilitarianism engages in no such rule-worship. It is, however, recognized to be a radical doctrine that implies that

many of our ordinary moral feelings may be mistaken. In this respect, it does what good philosophy always does—it challenges us to rethink matters that we have heretofore taken for granted.

If we consult what Smart calls our “common moral consciousness,” it seems that many considerations other than utility are morally important. But Smart is right to warn us that “common sense” cannot be trusted. That may turn out to be Utilitarianism’s greatest contribution. The deficiencies of moral common sense are obvious, once we think for only a moment. Many white people once felt that there is an important difference between whites and blacks, so that the interests of whites are somehow more important. Trusting the “common sense” of their day, they might have insisted that an adequate moral theory should accommodate this “fact.” Today, no one worth listening to would say such a thing, but who knows how many other irrational prejudices are still a part of our moral common sense? At the end of his classic study of race relations, *An American Dilemma* (1944), the Swedish sociologist Gunnar Myrdal reminds us:

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

Could it be, for example, that future generations will look back in disgust at the way affluent people in the 21st century enjoyed their comfortable lives while third-world children died of easily preventable diseases? Or at the way we slaughtered and ate helpless animals? If so, they might note that utilitarian philosophers of the day were criticized as simple-minded for advancing a moral theory that straightforwardly condemned such things.

2

THE GOOD

2.1 Promoting the Good

In this first half of the book, we are going to examine various factors that might be thought to possess moral relevance at the normative level. But the list of potential “candidate” factors is virtually unlimited; how should we decide which ones to discuss? My suggestion is quite simple. Most of us (the readers of this book, at any rate) share a common moral outlook which we might call *commonsense morality*. People may differ about the details, but at least the broad features are familiar and widely accepted. Even those people who reject commonsense morality—whether in whole or in part—are typically quite familiar with it (it may influence their moral intuitions, for example, even if these are intuitions that, on reflection, they are prepared to disavow). I implicitly appealed to commonsense morality in the example of the drowning woman (in 1.4), when I was trying to give intuitively plausible examples of morally relevant factors. I was counting on your acceptance of commonsense morality—or, at least, your familiarity with it. Since commonsense morality recognizes and endorses the four factors that I mentioned, I assumed that you would either accept these factors yourself or else recognize them as factors that many other people regard as relevant.

This feature seems to me to be a desirable one when selecting factors for more extended examination as well: since we will obviously have to limit our discussion, it makes sense to select factors that we will all recognize as being widely accepted. So I propose to let the common moral outlook continue to be our guide, at least so far as setting the agenda goes: the factors that we will discuss in Part One will be those that are recognized by commonsense morality. Note that we can do this without necessarily *endorsing* the views of commonsense morality. We are only letting commonsense morality set the agenda for discussion; we are not assuming that all of the factors it recognizes are of genuine moral relevance.

It is somewhat arbitrary where we begin, but for reasons that will emerge as we go along, I think the best factor to consider first is this: the goodness of outcomes. This is a factor that I think virtually everyone recognizes as morally relevant. It may not be the *only* factor that is important for determining the moral status of an act, but it is certainly at least *one* relevant factor. If an act will have bad results, that is a reason not to perform it; if, on the other hand, it will have good results, then that is a reason *to* perform it.

Sometimes, of course, no matter what we do, the outcome will be bad. Then all we can do is make the best of a bad situation. But this is still a matter of the goodness of outcomes. All other things being equal, we should pick the act where the outcome will be the least bad overall, for that will be the act with the *best* outcome (of those available to us). Similarly, and more happily, sometimes we have to choose between two acts *both* of which will have good outcomes. In such cases, all other things being equal, we should pick the act with the *better* of the two outcomes.

This talk of “all other things being equal” is just another way of recognizing that goodness of outcomes may not be the only morally relevant factor that comes into play in any given case. Even if a particular act would lead to better results overall than any other act available to the agent, perhaps there could still be *other* factors that oppose performing the given act, so that, on balance, the act is morally forbidden. If so, then in cases where this happens other things are *not* equal—and so “all things considered” the morally best act may not be the act with the best *results*. We won’t start asking until the next chapter when, if ever, the goodness of outcomes can be “outweighed” in this fashion. But the point to be emphasized, once more, is this: even if goodness of outcomes is not the only factor, it is certainly at least *one* morally relevant factor. So, all other things being equal, we should pick the act with the best results. We should pick the act that best promotes the good.

It is important to understand that in saying that the moral status of an act is determined (at least in part) by its results, this is meant to include *all* of its results. It is not only the immediate, or short term, results that matter: long term results, side effects, indirect consequences—all these matter as well, and they count just as much as short term or immediate consequences. If, for example, I must choose between an act with a small immediate positive effect—but no other later effects—and an alternative act that will have no immediate effect, but will eventually produce a lot of good, it is the second act I should perform. Similarly, if an act will have both good results and bad results, then these must *all* be taken into account. The question is: how good or bad will the results be overall, on balance, taking into account all of the results; and how does this compare to the overall results of the other acts available to the agent?

Now it was probably obvious all along that talk of “good outcomes” or “good results” was meant to include all of the results and all of the conse-

quences. But for all that, it must be noted that these terms are indeed being used here in a somewhat nonstandard fashion. For normally when we talk of the “results” or “outcomes” or “consequences” of an act, we mean to be restricting our attention to those events taking place *after* the act that are *caused* by the act. It is only if the act literally produces some later event that we call that event one of the act’s results or consequences. For example, if the later event would have happened anyway, we don’t normally consider it part of the consequences of the act. But for our purposes in normative ethics, we need to give these terms a “wider,” more inclusive sense than is normal. In thinking about the goodness of outcomes, what matters is not the goodness of what I literally produce myself, but rather the goodness of everything that ultimately happens, the goodness of the “upshot.”

An example should make this clear. Suppose five people are in danger of losing their lives, but if I try to rescue any of them at all, I will only be able to save one of them. Luckily, however, there is someone else willing to perform the rescue—and she will indeed be able to save all five, provided that I don’t get in the way. (If I am in the way, she won’t be able to save any.) Now what should I do? If I sit back and let the other person perform the rescue, I don’t actually produce any good at all myself, whereas if I do try to perform the rescue, my act would at least produce the good outcome of one life saved. So should we say that I ought to try to perform the rescue, since this is the act that will have better results, in the normal narrow sense of that term?

But this is clearly the wrong answer. Even though I will save one, the other four will die. Yet were I only to sit back and allow the woman to rescue all five, then no one would die at all! This second outcome is obviously the better one from the moral point of view, even though in terms of the narrow consequences of my act this may not be the best choice. What this shows, of course, is that in thinking about goodness of outcomes as a morally relevant factor, our real concern should be with the overall upshot—how good or bad things will be if I act in one way rather than another—rather than with what I myself produce. Accordingly, when we talk of the results, consequences, or outcome of an act, let us understand these terms in a wide sense, a sense that encompasses everything that happens, the entire upshot. (The example also shows, incidentally, that sometimes the act with the best results will be the “act” of doing nothing at all.)

So when we look at the results of an act we want to look at everything that happens. This stretches the normal notion of “results” somewhat, but seems to be necessary to get at the morally relevant concept. But we may need to stretch the notion of “results” in a second way as well: in principle we may need to look not only at what *will* happen (if the act is performed) but also at what *is* happening (when the act is performed). In some cases this may be a fairly obvious point. For example, a given act may itself be quite difficult to perform. In deciding whether that act has good results

overall, we will surely want to include, say, the physical pain the agent suffers *while* performing the act, and not only the good and bad results that may come after the act is completed.

Other examples will be more controversial, and will depend on the specific theories (yet to come) concerning exactly what features make an outcome good. On at least some views, however, the occurrence of an act of a given type may in itself be a good or bad thing that can happen; if so, then in assessing the goodness of outcomes we will want to include the agent's act itself as a potentially significant part of the "outcome" of the act.

Even if we stretch the term this far, it may seem obvious, at any rate, that there can be no need to take into account—as part of the act's "outcome"—what happens *before* the agent acts. After all, no act can alter the past, and so whatever good or bad events occurred in the past will be the same no matter what the agent does now; so the past cannot be relevant in determining whether one act has a better outcome than another. In fact, however, this too will depend on the details of one's theory of the good. On some views it will matter whether the past and the future "fit" together in certain ways (for example, we might think the outcome is better if certain people who have done evil in the past are punished in the future). On views like this, we might say that what matters is how well the course of world history goes. The upshot of some act might be that history as a whole goes less well, even though we wouldn't see this if we looked merely at what happens during or after the act. I think we should leave room for the possibility of such views, and so we will have to stretch the notions of "outcome," "results," and "consequences" so that they include absolutely everything that happens if an act is performed—past, present, and future.

This is certainly not to say that absolutely everything that happens *does* affect the goodness of the outcome. (Obviously, much, or most, of what happens is neither good nor bad in itself.) We still need an account of exactly which features go into making one outcome better than another. And this will, in fact, be our major concern in this chapter. But one final preliminary point still needs to be clarified. When we say that one morally relevant factor in determining the moral status of an act is whether that act will promote the good, we do not mean to be taking into account every kind of moral goodness that an act might produce. Our concern, rather, is only with *intrinsic* goodness.

Many things with moral value possess only *instrumental* value. They are valuable only insofar as they can contribute to producing other goods (or eliminating various bads). Thus, for example, medicine, air conditioners, automobiles, and money all possess instrumental value; they are useful means for acquiring or producing other goods (or eliminating bads); among other things, they can be used to produce various pleasures and to eliminate or avoid various pains. For most of us, these things are valuable *only* because of this instrumental value. On reflection we would agree that we do

not really value these things for their own sake, or in themselves. They are merely instrumental goods, not intrinsic goods. In contrast, most of us value pleasure for its own sake; we think of it as something good in itself—not merely as a means to getting something else (though it may be that as well). Similarly, we think of pain as something bad in itself, something to be avoided not merely because of what else it might produce, but for its own sake. That is to say, we take pleasure to be intrinsically good, and pain to be intrinsically bad.

In determining the moral status of an act, it is only the intrinsic value of the results that matters. It would seem foolish, for example, to prefer one of two outcomes that were similar in terms of all the intrinsic goodness that they contained, simply because the first contained a far larger amount of instrumental good. After all, instrumental goods matter only because they can produce intrinsic goods. Ultimately, then, all that really matters are the intrinsic goods (and the intrinsic evils). So insofar as an act's moral status depends on the goodness of its outcome, what it depends on is the intrinsic goodness of the outcome.

This is not to deny that for many practical purposes it is handy to evaluate outcomes in terms of the instrumental goods they may contain. I might, for example, reasonably choose between two acts on the basis of which will produce more money for me—knowing all the while that money is of merely instrumental value. But in such cases I also know and expect that the extra money can eventually be put to work to produce some extra good that I value for its own sake. That is, I reasonably take the short term difference in instrumental goods as evidence of a long term difference in intrinsic goods. Ultimately, however, what really justifies my choice is this long term difference in intrinsic goodness.

Of course, all of this still leaves us with the difficult task of deciding exactly what is of intrinsic value (or intrinsic disvalue). What makes one outcome intrinsically better than another? This is the question we will be exploring in the rest of this chapter.

2.2 *Well-Being*

In principle, the intrinsic goodness of an outcome might depend on several different types of factors. That is, just as the moral status of an act might depend on the interplay of several different morally relevant factors—one of which is the goodness of the act's outcome—the goodness of the outcome may itself depend on the interplay of several different subfactors. (And the subfactors may themselves depend on several further factors—the “subsubfactors.” Since such talk can quickly get clumsy, typically I will simply call them all “factors,” whatever their level. It is, at any rate, largely a matter of convenience and custom whether we view the “subfactors” as forming a genuinely unified single “factor” at the next higher level—or in-

stead view talk of the “factor” as merely being shorthand for referring to the group of subfactors.)

As we shall see, it is controversial exactly what factors go into making one outcome better than another. But once more there is a point on which virtually everyone agrees: one such factor is how well off people are. That is, all other things being equal, one outcome is better than another if people are better off in that outcome. And, all other things being equal, if people are less well off, if their lives are going less well, then the outcome is worse.

As before, talk of “other things being equal” is to remind us that this may not be the *only* factor that matters; but it is, at least, *one* of the factors that matters. The higher the level of well-being, all other things being equal, the better the outcome.

But what, exactly, is well-being? What is it for someone’s life to go better or worse? What does someone’s being better or worse off consist in?

It is important to understand what we are *not* asking in raising these questions. We are not asking what sorts of things may have a causal impact on well-being. Obviously enough, almost anything at all might, in the right circumstances, affect someone’s level of well-being. And certain things will have an effect in almost all circumstances: money, a job, and opportunities, for example, will almost always raise your level of well-being; and poverty, disease, and enemies will almost always lower your well-being. But these things are only a means to raising or lowering your level of well-being. They have instrumental value (or disvalue) because of the impact they can have on how well off you are. But they are not themselves what well-being (or ill-being) *consists* in. What we want to know, then, is what is it that *constitutes* being well off? (Not: what are the *means* to being well off?) We want to know: what are the “components” of well-being?

One important and popular answer is this: well-being consists in the presence of pleasure and the absence of pain. The more pleasure and the less pain that a life contains, the better the life. Increase the amount of pleasure that a person experiences (or decrease the amount of pain), and you make that person better off; increase the pain (or decrease the pleasure), and you make her worse off.

Now in point of fact, pretty much everyone believes at least this much: the presence of pleasure and the absence of pain is at least one component of well-being. (It is quite hard to deny this. The value of pleasure and the disvalue of pain seem virtually self-evident to anyone experiencing them.) But the particular view I have in mind makes a bolder claim; it holds that well-being consists *solely* in the presence of pleasure and the absence of pain. Pleasure and pain are the *only* elements that directly constitute how well off a person is.

We can put this same point another way. If we ask ourselves what are the intrinsic goods by virtue of which a life worth living *is* a life worth living, the answer—according to this view—is pleasure, and only pleasure. Simi-

larly, the only intrinsic evil by virtue of which a life is less worth living is pain. This view—that well-being consists solely in the presence of pleasure and the absence of pain—might be called *welfare hedonism*.

Actually, there are at least four different hedonistic views that need to be distinguished. *Psychological hedonism* is the claim that—as a matter of empirical, psychological fact—everyone's ultimate goal is to pursue their own pleasure and minimize their own pain. *Ethical hedonism* is the claim that—as a matter of normative ethics—what everyone *should* do morally is to try to maximize pleasure and minimize pain. From a logical standpoint, at least, since these two views make no claims about what has or lacks intrinsic value, they need to be distinguished from *value hedonism*, which is the claim that pleasure is the only intrinsic good and pain the only intrinsic evil. But even this last view needs to be distinguished from welfare hedonism, which limits itself to the somewhat more modest claim that *well-being* consists solely in the presence of pleasure and the absence of pain. (After all, one might accept welfare hedonism and yet still believe that there are other intrinsic goods that have nothing to do with well-being.)

Some people are attracted to all four hedonistic claims. But it is only welfare hedonism with which we are concerned here, since it is only welfare hedonism that purports to be and limits itself to being a theory about the nature of well-being. So, for our purposes, we can put these other versions of hedonism aside and use the term “hedonism” to refer to welfare hedonism.

Of course, the fact of the matter is that life is not one unending run of intense pleasure, unblemished by any pain, and so the hedonist needs to tell us how to rank lives that are mixtures of various sorts of pleasures and pains. However, the basic idea here is fairly simple and obvious: add up the pleasures and subtract the pains. The greater the “sum” (that is, the higher the net balance of pleasure minus pain), the better the life. But since not all pleasures are equally good, and not all pains are equally bad, the view needs to be made somewhat more complicated. Still, a solution is not beyond our reach. Intuitively, it seems, the longer a pleasure lasts, the better; and the more intense the pleasure is, the better as well. So by multiplying intensity times duration, we can calculate the quantity of pleasure a given pleasant experience contains. Similarly, we can calculate the quantity of pain by multiplying the intensity of the pain times the duration of the pain. By adding up the total quantity of pleasure and then subtracting from this the total quantity of pain, we arrive at a measure of how good a life someone has had as a whole—their overall level of well-being. (Different calculations, but along similar lines, will allow us to make judgments about how well off a person is at a time, or during a certain period, or how much a person's level of well-being would vary under different outcomes.)

Advocates of such *quantitative hedonism* need not believe that we can actually make the relevant calculations, at least not with very much precision. But they can still claim that the basic ideas expressed here capture real

and important insights into the nature of well-being: well-being is simply a matter of pleasure and pain; and the value or disvalue of pleasures and pains is simply a matter of their intensity and duration.

One important implication of quantitative hedonism is that quantity for quantity, all pleasures are equally valuable. Of course, people will differ from one another with regard to what sorts of things *give* them pleasure: you may derive great pleasure from listening to opera, while I find the experience painful; I may enjoy bird-watching, while you find it boring. From a practical standpoint, therefore, it will clearly be important to get our facts straight about what exactly gives pleasure to whom. But so long as we are indeed talking about equal quantities of pleasure, it doesn't matter where the pleasure comes from. Similarly, some things (such as great art) may be rich and fruitful sources of large quantities of intense and abiding pleasure, while other things (such as watching television) are at best sources of minimal, mild, and fleeting pleasure. So from a practical standpoint it may well be rational to invest more energy in contemplating art than in watching television. But so long as we are talking about equal quantities of pleasure, it doesn't matter where the pleasure comes from: a unit of art-based pleasure is of no greater value than a unit of television-based pleasure.

This, at any rate, is the view of quantitative hedonism. But it is at just this point that some think quantitative hedonism goes astray. Many find it very hard to believe that one kind of pleasure is no more valuable than another (unit for unit). After all, if *quantity* of pleasure is all that matters, this seems to lead to some pretty bizarre conclusions. Consider the life of a contented pig, wallowing in the mud: it might be filled with pleasure, a far greater quantity of pleasure than can be found in the life of a person. Yet who would want to trade places with the pig? Who thinks that the pig has a better life?

Pressed to defend our preference for a human life over a porcine existence, a natural suggestion is this: pigs might have a greater quantity of pleasure, but humans can have pleasures of a much more valuable *kind*. Unlike pigs, humans can experience love, friendship, and art; they can have the pleasures of discovery, creativity, and understanding. We seem drawn to the thought that such "spiritual" or "mental" pleasures are simply of a higher quality than the mere "physical" or "bodily" pleasures that are all that are available to the pig—and that these "higher" pleasures are of greater intrinsic value than the "lower" pleasures. If this is right, then when judging the level of well-being of some individual (whether human or non-human) we must take into account not only quantity but also *quality*. (Perhaps a new formula could achieve this by multiplying intensity times duration times a "quality rating"—but we need not pursue this question. We can also put aside, for the sake of simplicity, the question of whether quality can also affect the value of *pains*, as well as pleasures.)

Such *qualitative hedonism* will be resisted by advocates of quantitative hedonism, who may claim that the appeal of qualitative hedonism is based

on a simple mistake. As we have already observed, and as the quantitative hedonists themselves would stress, the sources of the so-called higher pleasures typically lead to far greater *quantities* of pleasure than do the sources of the so-called lower pleasures. So from a practical point of view, it typically makes sense to prefer the experiences of the higher pleasures—for these experiences will simply contain more pleasure! But this doesn't show that a given quantity of higher pleasure is any more valuable than the *same* quantity of lower pleasure. In and of itself, quality just doesn't make a difference to the intrinsic value of (a unit of) pleasure.

In reply, the qualitative hedonists will simply deny that they are making any such mistake. Admittedly, sources of higher pleasures typically yield far greater quantities of pleasure than do sources of lower pleasures; but even when we take this into account—even when we focus on similar quantities of pleasure—we still find ourselves attracted to the thought that a given quantity of higher pleasure is more valuable than the same quantity of lower pleasure.

Like many of the other controversies in normative ethics that will be discussed in this book, I will not try to settle here which side is correct or which side has the stronger argument. For reasons discussed earlier (in 1.3) it should not surprise us if intelligent and thoughtful individuals can sincerely disagree as to which view seems most plausible when all the arguments are taken into account. (It is also worth bearing in mind that here—as elsewhere—further arguments might be offered on either side.) Consequently, all one can do—on this question, as on the others that will follow—is to weigh the various arguments of each side and to decide for oneself.

Of course, even if we agree with the qualitative hedonist, that higher pleasures are more valuable than lower pleasures, we are still in need of some test to tell which of two pleasures is the more valuable (that is, higher in quality). The most plausible suggestion that has been made is this: ask the experts. That is, ask those who have experience with both sorts of pleasures. If those who have experience with both pleasure X and pleasure Y prefer to have X rather than Y (given a choice between equal quantities of either), then we can reasonably conclude that X is the more valuable pleasure.

It must be admitted that applying this test is not without its difficulties. As often as not, it may be impossible to find someone who genuinely has first-hand experience with both kinds of pleasures. How many lovers of opera really know what it is like to be totally absorbed watching game shows on television? And what human being has had the experience of being a pig wallowing in the mud? Still, the qualitative hedonist can admit these practical difficulties while insisting that the basic idea behind the test is correct: pleasures can differ as to their quality; and these differences in quality would be reflected in the preferences of those who had experienced them.

A remaining worry is this. Is qualitative hedonism genuinely a form of *hedonism*? The welfare hedonist claims that pleasure (and the absence of

pain) is all that matters intrinsically for well-being. If one holds that quantity of pleasure is not all that matters, but that quality matters too, has one abandoned the commitment to hedonism? People's intuitions on this question seem to differ, although it is not clear what turns on the answer other than the label. Even if the so-called qualitative hedonist is not genuinely a hedonist, the position she holds might still be the truth.

However, pushed this far, we might want to go farther. If the value of our various experiences can vary—depending on what *kind* of pleasures they are—perhaps other experiences may have value, too, even though they are not pleasures at all! Perhaps certain experiences have a high quality and contribute to our well-being, even though it would be stretching things to say they are pleasant. At the very least, it might be grossly misleading to focus on whatever pleasure they contain, as this might be quite irrelevant to their value.

Perhaps, then, the position we should actually accept is this: all that intrinsically matters for determining one's level of well-being are the various experiences one has—one's mental states. Some mental states are more valuable than others, and pleasure is presumably one such kind of valuable experience (or one aspect of an experience that may make it valuable). But it is not the only kind of valuable experience (or the only aspect of an experience that can give it value). In short, perhaps we should simply accept a *mental state* theory of well-being.

Strictly speaking, of course, hedonism is a version of a mental state theory (or perhaps: it is a family of mental state theories). It agrees that your mental states are what determine your level of well-being, and it holds a particular view as to which mental states have value (namely, the pleasant ones). But we are now considering the possibility of theories that move beyond hedonism in allowing other kinds of experiences to have value as well. Of course, any fully specified mental state theory would need to tell us exactly what kinds of experiences have what kind of value. And it should be noted that, other than hedonism, there have been relatively few attempts to do this. But even if one is uncertain as to exactly which kinds of experiences contribute to well-being, and by how much, one might be quite convinced as to the fundamental truth of mental statism: your mental states are the only things that directly determine your level of well-being.

Many people find the idea of mental statism quite compelling. It lies behind the common saying that "What you don't know can't hurt you." What else *could* constitute well-being other than having the right mental states? Yet mental state theories seem to leave something important out, for there are cases in which all mental state theories lead to answers that are difficult to accept.

Imagine a man who dies contented, thinking he has achieved everything he wanted in life: his wife and family love him, he is a respected member of the community, and he has founded a successful business. Or so he thinks.

In reality, however, he has been completely deceived: his wife cheated on him, his daughter and son were only nice to him so that they would be able to borrow the car, the other members of the community only pretended to respect him for the sake of the charitable contributions he sometimes made, and his business partner has been embezzling funds from the company, which will soon go bankrupt.

In thinking about this man's life, it is difficult to believe that it is all a life could be, that this life has gone about as well as a life could go. Yet this seems to be the very conclusion mental state theories must reach! For from the "inside"—looking only at the man's experiences—everything was perfect. We can imagine that the man's mental states were *exactly* the same as the ones he would have had if he had actually been loved and respected. So if mental states are all that matter, then—since this man got the mental states right—there is nothing missing from this man's life at all. It is a picture of a life that has gone well. But this seems quite an unacceptable thing to say about this life; it is surely not the kind of life we would want for ourselves. So mental state theories must be wrong.

A defender of the mental state approach might insist that the businessman's life must have gone well, since he was happy. But those who think something important was missing from this man's life need not be impressed by this answer. Some will deny that the man was genuinely happy: he *thought* he was, but in reality, he was not; perhaps the example shows that there is more to happiness than having the right sorts of mental states. Others will agree that the businessman was indeed happy; but they will argue that this only shows that there is more to well-being than happiness, since something is certainly less than ideal about the deceived man's life. In short, whatever we feel about the correct account of happiness, the example still seems to show that mental state theories are wrong: having the best kind of mental states is just not enough.

To be fair, some mental state theorists may not think that our deceived businessman had the very best kinds of experiences—perhaps it would be better to have experiences of adventure or discovery. But it seems we could compose a comparable story whichever mental states are best. Indeed, if we want, we can imagine an "experience machine" that electronically stimulates your brain, giving you whatever experiences you think most valuable. The effect is so perfect that it will feel ("from the inside") *exactly* as though you were climbing mountains, composing symphonies, or what have you. In reality, you are simply floating in a tank in the scientist's lab—but once you are hooked up, all memory of this will disappear, and you will have exactly the mental states endorsed by your favorite mental state theory.

Would you want to spend your life hooked up to an experience machine? If not, then it seems that mental state theories must be wrong—all of them. For life on the experience machine gets the mental states exactly right; so if something is missing from a life on the experience machine—if it is not the

best kind of life imaginable—then there must be more to well-being than having the right kind of experiences.

But what is missing? A natural response to these examples—the deceived businessman, the experience machine—is that these people don't really have what they *want*. They *think* they do, but they don't. The businessman wanted to be loved and respected. He thought he was—from the inside it *seemed* that he was, he had the same mental states as someone who was—but still he wasn't *really* loved and respected, even though that was what he *wanted*. Similarly, the person on the experience machine may, say, have the very same mental states as someone who is climbing a mountain, and climbing a mountain may be what he most wants to do. But since he obviously is not really climbing any mountains, he doesn't truly have what he wants—to climb mountains! In short, what is missing from all of these examples is that the person's preferences are not actually satisfied; things are not in fact the way they want them to be.

The point could be put this way: what we want out of life is to have what we *want* out of life, and what we want is almost always far more than merely having certain kinds of experiences. For example, if you want to be a great and famous novelist, then your desire is only satisfied if you have in fact written a great novel, and your work is well-known by others. Merely having the experiences is not enough to actually satisfy the desire. (It should be noted that "satisfy" is being used here in a somewhat technical sense: a desire is satisfied, in this sense, if and only if the very thing that is desired actually occurs. This concept should not be confused with the *feeling* of satisfaction that often accompanies our believing that our desires have been met.)

These thoughts suggest a desire-based or a *preference* theory of well-being. According to the preference theory, well-being consists in having one's preferences satisfied. To the extent that your preferences or desires are satisfied, you are better off; to the extent that your preferences or desires are not satisfied, you are less well off. (Of course not all of your preferences are equally important to you, so we would probably want to make the view more complicated, holding that the stronger the desire, the greater the contribution to well-being made by the satisfaction of that desire. But we won't pursue these details here.)

Not surprisingly, mental state theorists will argue that the attractiveness of the preference theory is based on a confusion. Unlike our make-believe examples, they will note, in real life there is actually only one way to achieve the very same mental states one would have if the corresponding "external" states of affairs really existed—and that is to actually bring about the external states of affairs. For example, the only way to have the very same experiences as someone climbing a mountain is to actually climb a mountain. So from a practical standpoint, we are justified in simply aiming at the *external* states, knowing that, if we can achieve these, the mental

states will follow. Thus, aiming at the external states has instrumental value; it is the best means to the mental states. But this should not mislead us into thinking that the external states have *intrinsic* value, making a direct contribution to well-being; only the mental states themselves do that.

Preference theorists, however, may not be impressed by this argument. They will, of course, admit that in real life a given set of experiences and the corresponding external states of affairs go pretty much hand in hand—we don't get the one without the other. But this doesn't show that it is the former and not the latter that matters. We may have to think about various kinds of make-believe examples to find cases where we would have the mental states without the corresponding external reality; but when we do think about such cases, we see quite clearly that what we value is *more* than merely having the right mental states. Our preferences concern external reality, too, and not only experiences. Of course, it is certainly true that *one* of the things we all have preferences about is having the right mental states. (I want to experience pleasures and not pain; I want to have the experiences of loving and being loved; I want to experience the accomplishment of my various goals; and so on.) So a preference theory will certainly recognize that my mental states are one of the things that help determine my level of well-being. But mental state theories go wrong in claiming that mental states are the only thing that directly determine my level of well-being. My other preferences matter too; and they must be satisfied as well if I am to be completely well off.

Once more, I won't try to settle this dispute. But I do want to consider two other possible objections to the preference theory. Both try to show that the theory yields unacceptable results. The first objection notes that, according to the preference theory, I am made better off by the satisfaction of my various desires, regardless of the subject matter of the given desire. Suppose, then, that I am a fan of large prime numbers, and so I hope and desire that the total number of atoms in the universe is a prime. Imagine, furthermore, that the total number of atoms is, in point of fact, prime. Since this desire is satisfied, the preference theory must say that I am better off for it, that my life is the better for it. (I don't *know* that my desire is satisfied, of course. But as already noted, the question of whether a desire is satisfied should not be confused with the question of whether one realizes it and consequently has a feeling of satisfaction.)

But this is absurd! The number of atoms in the universe has nothing at all to do with the quality of my life. It makes no difference at all to my level of well-being. (Were the number really even, this would not leave me *worse* off!) So the preference theory must be false.

One possible explanation of what has gone wrong is this. My preferences can range over anything at all—including things that have nothing to do with me or my life. But the brute satisfaction of *these* preferences cannot plausibly be claimed to make me better off; so an *unrestricted* preference

theory is unacceptable. The only plausible approach will be in terms of a *restricted* preference theory: we will have to find the particular subset of my desires that concern my life, and claim that it is only the satisfaction of *those* desires that constitutes my well-being.

(The mental state theorist, of course, will insist that this misdiagnoses the problem. The real source of the absurdity is in thinking that the mere satisfaction of a desire—any desire—can, in and of itself, make me better off, even if I never know about it, and so even if it never has any effect on my mental states. But the preference theorist may reply that she has already admitted that one of the things we care about are our mental states, in particular knowing that our desires have been satisfied. So the preference theorist can acknowledge that if one of my desires is satisfied without my knowing about it, I am still not as well off as I might have been—since my desire to *know* about it is not satisfied.)

The second objection to the preference theory—even in its restricted form—is this. Many of my desires, including desires about my life, are based on misinformation, sloppy thinking, inexperience, prejudice, bias, and various other kinds of error or irrationality. I might want something, for example, because I am misinformed as to its nature, and I falsely believe I will enjoy having it. It seems implausible to claim that giving me what I want in a case like this makes me better off. (Well, strictly speaking, knowing that I had gotten my way might give me some mild feeling of satisfaction, and this would presumably make me slightly better off; but what seems implausible is the claim that satisfying this mistaken desire leaves me better off in and of itself.) Yet I do in fact have such mistaken (or crazy, or irrational) desires among my desires, and so the preference theory apparently has to say that satisfying these desires does indeed improve my level of well-being.

To avoid this result, we need to distinguish between *actual* preference theories and *ideal* preference theories. As the name suggests, actual preference theories claim that well-being is a matter of the satisfaction of your actual desires—the desires you do in fact have, whatever their basis. As we have just seen, this has the counterintuitive implication that satisfying even your irrational or misinformed desires leaves you better off. In contrast, ideal preference theories claim that well-being is a matter of the satisfaction of your “ideal” desires—that is, the desires you *would* have if you were fully informed, thinking clearly, free from prejudice and bias, and so on. Since you would not have misinformed, crazy, or irrational desires under such ideal conditions, they are not among your ideal desires; so, on this view, satisfaction of such desires does not contribute to your well-being. It seems, then, that preference theorists who want to avoid the second objection should embrace an *ideal* preference theory. (Of course, assuming that they want to avoid the first objection as well, they will actually need to embrace a *restricted* ideal preference theory.)

But once we have made the move to an ideal preference theory, we may want to move even further. After all, what exactly is it that explains why it is only ideal desires whose satisfaction will contribute to well-being? Why does having what we want leave us better off in those cases—but only in those cases—where the desires in question are ideal desires? We cannot say: having what we want makes us better off by virtue of the very fact that these are the things we want. For this view actually supports the *actual* preference theory, rather than the ideal preference theory. (The things we want include everything that we *do* want—not just those things we would want if informed; and some of the things we *would* want if we had ideal preferences we do not in fact already want.)

A tempting thought is this. Having the things we would want—if only we were fully informed, rational, free from bias, and so forth—is valuable, because these are the things that are truly worth having! That is, if we were fully informed, and so on, we would be in an ideal position to *recognize* which things have value and which things do not. And then our preferences would follow accordingly: we would prefer to have what has more value over what has less.

But this means that the appeal to preferences is not doing any actual work in explaining the source of the value. It is not that the various goods that we would want were we ideally informed are valuable because we would desire them; rather, we would want these goods because we would be able to *see* that they are indeed valuable, *independently* of our wanting them.

Thus we are led to *objective* theories of well-being, theories that hold that being well off is a matter of having certain goods in one's life, goods that are simply worth having, objectively speaking. Similarly, there may be certain objective bads or evils, the having of which simply leaves one worse off. Possession of the relevant goods and the absence of the relevant evils is what constitutes well-being. And the goods and evils themselves have intrinsic value or disvalue independently of our desires (actual or ideal); indeed, they have the particular value they have regardless of whether anyone is in a position to realize this.

Of course, any fully specified objective theory will need to provide us with a list of the various objective goods and evils that together determine one's level of well-being. Different objective theories will provide different lists. Indeed, it is possible to view hedonism as a kind of objective theory—one with a very short list! (Pleasure is the only objective good; pain the only objective evil.) But most objective theories will go beyond hedonism in offering longer lists—lists that include other mental states, and indeed include various external states of affairs as well (that is, they will include more than mental states alone). Typical goods on such lists might include accomplishment, creativity, health, knowledge, friendship, freedom, fame, and respect.

The preference theorist will hasten to point out that goods like these are typically desired by us (and certainly would be desired under ideal condi-

tions), so the preference theory too can recognize the contribution these things make to well-being. And the hedonist will hasten to point out that goods like these are typically of crucial instrumental value in achieving the most pleasant mental states, so hedonism too will recognize the contribution these things make to well-being. The objective theorist, however, thinks that both these approaches miss the mark: the hedonist goes wrong in failing to recognize that these goods have intrinsic value, and not merely instrumental value; the preference theorist goes wrong in failing to see that the intrinsic value these goods possess is had objectively—it is not based on or somehow derived from the fact that we do or would desire them.

Yet if desire is not the basis of the value of these goods, what is? And what is it, if anything, that unifies the various items on the list? Why do these things, and only these things, have objective value? Not surprisingly, answers to these questions are difficult to come by, and different objective theorists offer different answers. Many proposals take the form of *perfectionism*: the objective goods are those that are elements in an ideal or perfect human life, one which fully realizes the distinctive and essential characteristics of human nature. Of course, even within a broadly perfectionist approach there is room for considerable disagreement, both with regard to what constitutes human nature and—accordingly—what goods are elements in the perfect human life.

But however we settle these questions, a difficulty remains. As we have seen, the objective theorist holds that possession of the objective goods makes one better off—regardless of whether or not one realizes this. This seems to have the implication that your life could be made better off by the possession of some “good” even though you yourself dislike it and would greatly prefer to be without it: since the good possesses objective value, your own opinion on the subject is quite irrelevant. Your life could be going well even though you are unhappy with almost all its central features! Thus, like the other theories of well-being, the objective theory too seems to lead to unacceptable results.

Some objective theorists may be willing to live with this implication of their theory. Others, however, may try to resist it, by claiming that although the value of the objective goods is not *based* on one’s preferences or pleasure, nonetheless it is true for each objective good that if one possessed it, one would be glad to have it. But it is hard to see what could guarantee the truth of this fortunate convergence of objective value and desire or enjoyment. Indeed, critics of objective theories may well think that there simply *is* no way to guarantee this result: if we want convergence of goodness and desire or enjoyment, then we must build desire or enjoyment right into the definition of goodness. That is to say, we must appeal to either hedonism or a preference theory.

This then brings us full circle. Each theory of well-being has its attractive features, but none is without its drawbacks. Although there is considerable

agreement that well-being is a central factor in determining the goodness of an outcome, there is nothing approaching agreement as to what exactly well-being consists in. As always, this does not mean that no single view is correct; it does, however, mean that each of us will have to decide for himself or herself which view is the most plausible one on balance—once all of the arguments of each side have been considered.

It is worth noting, finally, that despite this philosophical disagreement as to the precise theoretical basis of well-being, there is in fact a considerable consensus concerning most real-life cases. Advocates of the various theories may differ as to whether various goods are merely instrumental to well-being or are instead directly constitutive of well-being. But when considering realistic cases, we find that most plausible views are in a fair amount of agreement as to who is well off and who is not. So from a practical standpoint, at least, our inability to resolve the theoretical dispute may not be debilitating.

2.3 The Total View

Let's suppose that you've come away from the previous section with an adequate theory of well-being. It will be a theory of *individual* well-being, that is, a theory that tells us what it is for a given individual to be better or worse off. Such a theory is obviously useful for us, given our belief that well-being is at least one factor that determines the overall goodness of outcomes. But we can't use it as it stands, till we decide what to do about the fact that there is more than one person in the world. (There are, of course, millions and billions of people in the world; but the logical problem arises as soon as there are two!) For each person—let us suppose—we can form a judgment concerning how well off that person is and how that person's level of well-being would vary under different outcomes. But what do we do with all of this information? Whose well-being are we to take into account, in deciding which outcome is better?

One possible view is this: each one of us should take into account only his or her own level of well-being. If this view is correct, then when I say that a given outcome is better, what I should mean is that it is better in terms of my own well-being. And when you say that a given outcome is better, what you should mean is that it is better in terms of *your* well-being.

This view doesn't deny that one person's well-being can be affected in various ways by the level of well-being of another. Nor need it deny that the outcome that is best for one person might well be best for many others as well. It is just that when any given person evaluates an outcome in terms of well-being, all that *counts*—directly, in and of itself—is that particular person's level of well-being. Someone's saying that the outcome is good is just a way of saying that it is good *for them*.

A different possible view holds that each one of us should take into account *everyone's* level of well-being. When I say that a given outcome is

better, then, I mean that it is better counting everyone—not just me. And when you say that a given outcome is better, you mean the same thing: better, counting everyone.

Of course, this second view would be very similar to the first, if each of us counted everyone, but only for a very small amount, while counting ourselves for a great deal. So let's add, straight off, that according to this second view each of us should not only count everyone but should count everyone *equally*.

Now we agreed at the start of this chapter that one factor that determines the moral status of an act is the goodness of the outcome of that act. And we also agreed that well-being is at least one subfactor that helps determine the goodness of an outcome. But now we see that we have to decide how this factor—goodness of outcome—is to be understood: is the morally relevant factor only how good the outcome is *for me*, or is the morally relevant factor how good the outcome is *overall*, taking *everyone's* well-being into account?

So long as we clearly bear in mind that what we are asking is what matters from the *moral* point of view, I think that most of us will agree that everyone's well-being needs to be taken into account. If an act will benefit me, this is certainly a good result; but if it will benefit others as well, this is an even better result, morally speaking. Indeed, an act might have good results, from the moral point of view, even if there is no benefit in it for me at all. I pretty much took all this for granted previously: for example, in discussing the case of the drowning woman (in 1.4), I assumed we would all agree that saving the woman's life was a good result. And I assumed I could say this even without knowing whether my own level of well-being would somehow be improved by my act as well. So when we are considering goodness of outcomes as a factor that can help determine the moral status of an act, everyone's well-being counts.

(What could tempt us to think otherwise? Well, we might get the question of what counts, morally speaking, confused with the question of what counts from the *rational* point of view—or we might think the two must coincide (see 1.2). And then we might also hold the view that the only thing that I can count from the rational point of view is my own well-being. But if you do indeed accept this last claim about rationality—and I for one see no reason to believe it—then it seems much more plausible to *deny* that the rational point of view and the moral point of view must coincide. For it is extremely implausible to suggest that *morally* speaking, no one counts but myself.)

In talking about goodness of outcomes as a morally relevant factor, then, we are not just talking about goodness from a “personal” point of view or a “subjective” point of view—what's good for some particular individual. We are, rather, talking about goodness from what is sometimes called an “objective” point of view, an “impartial” point of view, or an “impersonal” point of view. (None of these labels is perfect, and the last is particularly

unfortunate: talk of what is good from the “impersonal” point of view might suggest that we are considering a standpoint that has nothing to do with persons! Better labels might be that of the “transpersonal” or “omnipersonal” point of view. But use of “impersonal” in this context is pretty well established.)

Suppose we agree then that goodness of outcomes is a matter of the overall objective good, the goodness of the results taking everyone’s well-being into account. This immediately raises a new problem, given that what is good for one person may not be good for another.

Imagine a very simple case. Suppose I must choose between producing outcome A and outcome B, and my choice will only affect two people. If I pick A, Anatole will be better off than if I pick B; but if I pick B, Belinda will be better off than if I pick A. So outcome A is better for Anatole, and B is better for Belinda. But which outcome is better taking *both* of these people into account? (Strictly, we should have asked: which outcome is better taking *everyone* into account? But since no one else is affected, these seem to come to the same thing in this case.)

Obviously, what would be best of all would be an outcome in which both Anatole *and* Belinda are benefited. Unfortunately, however, this is impossible, and only *one* person can be benefited. But who should it be? Is the outcome better if Belinda is benefited, or is it better if Anatole is benefited?

A natural proposal is this: it depends on who will benefit *more*. For example, if Anatole will be much better off in A than in B, while Belinda would only be slightly better off in B than in A, then A is the better outcome overall: admittedly, Belinda is the “loser” in A, but she does not lose very much; while if we picked B, then *Anatole* would be the loser, and would lose much more (that is much more than what Belinda loses in having A rather than B).

Now consider a slightly more complicated case. Suppose that Beatrice will be affected by the choice too. Like Belinda, Beatrice is better off in B than in A. Finally, suppose that although it is still true that Anatole stands to benefit more than does either Belinda or Beatrice individually, this isn’t true by very much. When we tell the story this way, the natural suggestion seems to be that B is the better outcome. Admittedly, the benefit to Belinda isn’t as big as the benefit to Anatole would have been, and the benefit to Beatrice isn’t as big as the benefit to Anatole would have been; but when you consider the benefits to both Belinda *and* Beatrice, this just seems to outweigh the “loss” to Anatole. Intuitively, that is, the individual benefits to Belinda and to Beatrice add up to a greater “total” benefit than we would get by benefiting Anatole by himself.

Generalizing, then, we might say that one outcome is better than another—at least so far as well-being is concerned—if the total benefit of having that outcome rather than the other is greater. But if we do state the point this way, we have to be careful not to lose sight of the basic, under-

lying idea that how good an outcome is depends on *everyone's* level of well-being in that outcome. That is, it's not just a matter of those who happen to get an extra *benefit* from that outcome.

Not surprisingly, we can often make comparisons between outcomes more easily if we put everyone else aside and focus only on those whose level of well-being varies from one outcome to the next; indeed, as we have just seen, we can often make the comparisons more easily still if we focus only on the *differences* in the levels of well-being (that is, the extra benefits). But for all that, the underlying idea really seems to be this: how good an outcome is depends on the *total* amount of well-being in that outcome. Ultimately, that is, the goodness of an outcome is not just a matter of the extra *boosts* to well-being that an outcome happens to contain (compared to the others); it is really a matter of *all* the well-being that it contains. All other things being equal, then, the greater the total amount of well-being, the better the outcome. Call this the *total view*.

It is important to avoid a common misunderstanding about this view. It does not say that goodness of an outcome is a matter of *how many* people are better off in that outcome; it is not a matter of "majority rules." Often, of course, the outcome with the greatest total amount of well-being will indeed be the outcome that benefits the greatest number of people. This is true, for example, in the case of Anatole, Belinda, and Beatrice. But it is also possible to have cases where this is not true. Suppose that Belinda and Beatrice would actually only be very slightly better off in B than in A, while Anatole will be *much* better off in A than in B. In this case, we can imagine, the total benefit to Belinda and Beatrice of having B rather than A is still less than the benefit to Anatole of having A rather than B; that is, even though only one person would benefit from A, while two would benefit from B, it is A that has a greater total amount of well-being. According to the total view, then, this is the better outcome.

It should be noted that all this talk about comparing who gets benefited more presupposes that it makes sense to talk about how *much* someone benefits, and whether this is more or less than someone else. Similarly, talk of adding up the total amount of well-being in a given outcome presupposes that it makes sense to talk about *measuring* a person's level of well-being—giving it a number that we can meaningfully compare to someone else's number—and then adding these numbers up. That is, this entire approach presupposes that it makes sense to talk about *interpersonal comparisons of well-being*. (Since another term for "well-being" is "utility," these are also often called "interpersonal comparisons of utility.") Yet it is often thought that there is something deeply problematic about such comparisons. It is sometimes suggested that such comparisons are simply meaningless, or perhaps meaningful but in principle impossible to make. Other times it is suggested that although in theory it may be possible to make the comparisons, in actual practice they cannot be made, or cannot be made with any kind of precision.

There are obviously a number of different objections being raised here (although they are often run together), and we do not have the space to try to answer them adequately. One point that does seem worth making, however, is that in evaluating the plausibility of a given objection we do well to keep in mind the details of our particular theory of well-being. (I see no reason to assume that interpersonal comparisons of well-being will be equally problematic regardless of what theory of well-being we hold.) For example, suppose we are hedonists. Then to claim that interpersonal comparisons are simply meaningless is equivalent to saying things like this: it makes no sense to say that one person is in more pain than another. This seems quite implausible. On the other hand, it does seem plausible to suggest that we often cannot compare the intensity of two people's pains with any great precision.

Perhaps this last point should be conceded more generally. In real life we are often unable to make anything more than rough and imprecise comparisons. But we certainly do seem to be able to do at least that much in many cases. And often our judgments, although rough, are sufficient to enable us to say that one outcome is better than another. So I am inclined to think that any general skepticism about the possibility of making interpersonal comparisons of well-being is probably misguided. On the other hand, it must be admitted that in most cases the assignment of precise numbers will probably be beyond us.

Of course, even if for all practical purposes precise numbers are not available to us, there still may be a fact of the matter as to exactly how much better off one person is than another. (In practice, all sorts of measurements may elude us, even though there is still a fact of the matter as to precisely what some quantity comes to.) Alternatively, it may be that the fact of the matter is that, even in principle, interpersonal comparisons of well-being can only be rough and imprecise. Even if this is so, however, the total view may still be correct: perhaps the total amount of well-being contained in an outcome can only be specified roughly and imprecisely; but for all that, it might still be true that when one outcome does contain a greater total amount of well-being than another, it is the better outcome.

In what follows, I will occasionally assign numbers when comparing levels of well-being. This will help to make the examples clear, but nothing should turn on the precise numbers used, nor on the numbers being precise. Consequently, even those who think that interpersonal comparisons of well-being cannot be precise—not even in principle—need not reject, for that reason alone, the arguments to come. (However, those who dismiss the possibility of interpersonal comparisons altogether will definitely need to reject a certain amount of what follows.)

Difficulties of measurement aside, there are two other issues concerning the total view that should be mentioned, even if only briefly. (They are too complex to pursue in any detail here.) The first is this. According to the to-

tal view, the greater the total amount of well-being, the better the outcome (all other things being equal). But there are two ways that we can increase the total: (1) we can make the people who actually exist better off; or (2) we can increase the *number* of people who exist, adding extra happy people. Now virtually everyone agrees that the first method does indeed make the outcome better: if you have a group of people, then the better off they are, the better the outcome. But it is far less clear whether the *second* method of increasing the total actually makes for a better outcome. Does the mere fact that an outcome contains *extra* happy people make the outcome better, in and of itself? Suppose the first outcome contains one billion people, all at a high level of well-being, while the second outcome contains one billion and *one* people, all at that same high level. The “original” people are no better off in the second outcome—does the mere addition of the *extra* person make the outcome better?

Many people comfortably answer yes. Such people can continue to accept the total view, since this is the answer it supports. Others, however, are inclined to think that the mere addition of extra happy people does not make an outcome better. Accordingly, such people must reject the total view, and find an alternative account of the contribution that well-being makes to the goodness of an outcome.

One common proposal is the *average view*—which holds that the greater the *average* level of well-being in an outcome, the better. (To get the average, take the total amount of well-being, and divide by the number of people in the outcome.) Now in those cases where the number of people is the same in both outcomes, the average view will still agree with the answers given by the total view. But in cases where one outcome has more people than another, the two views may differ. Since adding the extra happy people may increase the total without increasing the average, the total view will have to say the second outcome is better, while the average view can deny this. To return to our example, adding the one extra person to the original billion people will increase the total amount of well-being, but it does nothing to increase the average level of well-being. So the average view can hold that the second outcome is no better than the first.

There are, however, various problems with the average view. First of all, if the extra happy person is at a *higher* level of well-being than the original people, then the addition of the extra person may actually *increase* the average. So the average view will have to hold that the outcome with this extra, above-average person is better. And while some may find this intuitively acceptable—certainly friends of the *total* view will find this congenial—others may feel that the average view here fails to capture the original intuition that the mere addition of *extra* people cannot make for a better outcome. Furthermore, if the extra person is *below* average—even if only by a very small amount—then the second outcome will have a *lower* average, and so the average view will have to hold that it makes things

worse to have this extra person. But many will find it hard to believe that the mere addition of an extra happy person can make an outcome worse. (It may not make things better; but that is a far cry from saying it makes things worse.)

Thus those who are dissatisfied with the total view in its application to "different number cases" may find the average view unacceptable as well. In point of fact, finding an acceptable principle to govern different number cases may turn out to be a surprisingly difficult undertaking. In what follows, however, I will—if only for the sake of simplicity—assume the truth of the *total* view. Luckily, this choice should affect little of the ensuing discussion.

The final question about the total view concerns the well-being of *animals* (or, more precisely, nonhuman animals). When comparing the goodness of outcomes, does the well-being of animals count too?

Most people will readily agree that animals count too. If one can eliminate, say, the pain and suffering of a dog or even a bird, this makes the outcome better. But do animals count as *much* as people? If a cat's well-being is increased by a certain amount, does this make the same contribution to the goodness of the outcome as would a comparable increase in the well-being of a *person*?

Many think not. They hold that although animals count, they count for less than people. The well-being of animals matters, but it gets less *weight* than does human well-being. Such a view could get quite complex, assigning different weights to different *kinds* of animals: for example, perhaps highly intelligent animals like apes and dolphins count for a great deal; dogs and cats somewhat less; birds and fish even less; and insects not at all. Clearly, adopting a position of this sort would require a certain amount of revision in the total view (which, as it stands, treats all increases in well-being equally).

It seems possible, however, that our inclination to count the well-being of animals for less than that of humans is instead simply a matter of prejudice and confusion. Of course, in many cases the welfare of an animal might not be affected as much as the welfare of a human would be: slapping a cow, for example, might not cause as much pain as would slapping a human child. But if we are genuinely comparing the same *amount* of pain, it is hard to see why the pain of a cow should count for less than the pain of a human: after all, the cow is, by hypothesis, in just as much pain. Perhaps, then, unit for unit, the well-being of animals should count just as much as the well-being of people. And if this is the position we take, we can continue to accept the (unrevised) total view. (Holding this view, it should be noted, is fully compatible with claiming that some of the most valuable components of well-being—those elements that make the greatest contribution to individual well-being—may not be available to animals. For example, animals cannot enjoy poetry. Once again, however, this does not give us any reason to count animals less when we are talking about *comparable* elements of well-being.)

The moral status of animals is a difficult and controversial topic. What's more, it is obviously of considerable practical significance. But we will not consider it further here. Indeed, except for one or two brief remarks, the special moral issues that arise for animals will not be considered again. I want to emphasize, however, that this omission is a response to the complexity of the topic, not a claim about its unimportance.

2.4 Equality

Is the total amount of well-being the only factor that affects the goodness of an outcome? Many people think not. For if the total is all that matters, then it doesn't matter how the well-being is *distributed*. So long as the total amount of well-being is the same, it doesn't matter who receives how much.

Suppose we have two possible outcomes—each with ten people, and each with a total of 1,000 units of well-being. (Units of well-being are sometimes called “utils”; hedonists often favor the term “hedons.”) In the first outcome, each of the ten people has 100 units. In contrast, in the second outcome, nine of the people have only 50 units each, while the tenth person has all of the remaining 550 units!

As far as the total view is concerned, these two outcomes are equally good—since the total amount of well-being is the same in both. Yet almost everyone is drawn (at least initially) to the thought that the first outcome is far better than the second. For in the first outcome the well-being is distributed *equally*; everyone is at the same level of well-being. But in the second outcome, there is a huge difference between the level of the one well off individual and everyone else. And this inequality seems to make the second outcome worse.

Cases like this seem to support the conclusion that the total amount of well-being is not all that matters intrinsically in determining the goodness of an outcome. Distribution of the well-being matters too, and in particular, equality matters.

Now this conclusion is not necessarily inconsistent with the total view. One could understand the total view as merely holding that the total amount of well-being is *one* factor that can contribute to the goodness of an outcome, so that—all other things being equal—the greater the total, the better the outcome. One could believe this and yet still hold that the total amount of well-being is not the *only* factor that matters intrinsically: equality (or, more generally, distribution) matters too.

Obviously, however, the view that equality matters *is* opposed to any view that says that *all* that matters is well-being. Let us call a view of this bolder type—one that claims that well-being is *all* that matters—*welfarism*. And let us call the view that says that equality matters too, *egalitarianism*. What most of us believe, then, is that welfarism is mistaken: distribution matters too, and in particular, equality matters. We are not welfarists, but egalitarians.

(A word about this terminology might be in order. I have defined the welfarist as holding the “extreme” or “pure” view that welfare is the *only* factor that matters, whereas—in contrast—I have defined the egalitarian as merely holding the “moderate” view that equality is *one* of the factors that matters. The reason for the nonsymmetrical treatment of these terms is that virtually everyone accepts at least “moderate” welfarism, and virtually no one accepts “pure” egalitarianism; so the interesting question concerns the choice between pure welfarism and moderate egalitarianism. For simplicity, therefore, I’m using the unadorned, unqualified labels “welfarism” and “egalitarianism” to refer to the two positions at the heart of the controversy.)

It might be thought that the welfarist is inconsistent in not admitting that equality of distribution is one factor that helps to determine the goodness of outcomes. After all, when explaining and defending the total view we observed that from the moral standpoint it seems that everyone counts equally. (That’s why we have to add together the total of *everyone’s* well-being, counting everyone equally.) It seems, then, that the total view recognizes, and indeed presupposes, the importance of equality. So how can it be denied now?

There are, however, two importantly different notions of equality at issue, and an advocate of the total view can consistently endorse the importance of one while rejecting the other. The total view holds that people are equal in this sense: in and of itself, each unit of well-being makes an equal contribution to the overall goodness of an outcome, regardless of *whose* well-being it is. (The welfare of blacks counts as much as the welfare of whites, an increase in the well-being of a woman counts as much as a comparable increase in the well-being of a man, my welfare counts neither more nor less than yours, and so on.) But accepting this does not entail that it matters in and of itself whether everyone is at an equal level of well-being. The equal *importance* of well-being simply does not entail the importance of equal *levels* of well-being.

Thus it is not true that the welfarist is somehow being inconsistent in not recognizing the intrinsic value of equal distribution. Welfarism can consistently embrace one claim about equality (people matter equally) without embracing the other claim (equal distribution matters).

By the same token, however, this also means that the welfarist cannot attempt to dismiss egalitarianism on the grounds that welfarism has *already* taken the significance of equality into account. For the egalitarian can correctly observe that it is only one kind of equality that matters for welfarism, whereas the egalitarian thinks that *both* kinds of equality are important.

So the question remains: does equality of distribution matter? Is equality—that is, equality in the distribution of well-being—one of the factors that determines the goodness of an outcome? (For simplicity, we can now revert to the practice of letting “equality” be shorthand for “equality of distribution,” since that will be our focus for the rest of this section.)

As we have seen, most people do find themselves intuitively drawn to the thought that equality does matter. But the welfarist argues that belief in egalitarianism is due to a natural mistake. To understand it, the first thing to bear in mind is that in real life we cannot directly distribute well-being itself; all that we can do is to distribute resources—the *means* to well-being—hoping thereby to produce well-being indirectly. The second thing to note is that so far as the distribution of *resources* is concerned, equality does indeed matter, as a way of maximizing the total amount of well-being produced; that is, it matters *instrumentally*. From a practical standpoint, then, it will often make sense to aim at equal distribution. This tempts us (says the welfarist) into thinking that equal distribution matters in and of itself—and so equal distribution of well-being must matter as well. But the egalitarian is wrong to ascribe intrinsic value to the equal distribution of well-being. First of all, it is only equality of resources that matters—not equality of well-being. Secondly, equality matters at best only instrumentally; it has no intrinsic value.

The welfarist's argument is a difficult one, so let's consider it at greater length. The heart of the argument turns on the observation that equality often has *instrumental* value; more precisely, equality often has instrumental value when we are concerned with the distribution of *resources*. This is because most resources, for example money, display "diminishing marginal utility": the more money you have, the smaller the increase to your well-being that will be produced by an extra dollar. This means that we can typically produce the greatest total amount of well-being by distributing resources equally.

Suppose, for example, that we have two people—a man and a woman—and a million dollars to be distributed between them. One possible distribution would be to give the woman the entire million dollars, leaving the man with nothing. Here, presumably, the woman will be at a high level of well-being, while the man will be at a low level. Alternatively, we could give the man one thousand dollars, leaving the woman the million less the thousand. When compared to the first outcome, the level of well-being of the poor man will increase dramatically; in contrast, the wealthy woman's level of well-being will go down only slightly. (When you are poor, an extra thousand dollars produces a lot of well-being; when you are wealthy, it may barely matter.) So the *total* amount of well-being produced by the money will be greater in the second scenario. And if we took another thousand dollars and "shifted" it as well, we would have an even *greater* total amount of well-being. (Perhaps the poor man would not gain quite as much extra well-being this time as from the first thousand, but it would still be much greater than what the wealthy woman would lose.) All other things being equal, then, we would keep getting a greater and greater total "bang for our bucks" the closer we came to having the money distributed equally. We could expect to produce the greatest total amount of well-being in the

scenario in which the man and the woman each have half the money. That is to say, equal distribution of money may be instrumentally valuable as a way of achieving the greatest total amount of well-being.

This may not be true in every single case, but it seems true often enough: equal distribution of resources helps to achieve the highest possible total amount of well-being. Reasonably enough, then (the welfarist continues), we find ourselves drawn to the thought that equality matters. But it is a mistake—albeit a natural one—to take this to mean that there is any *intrinsic* value in the equal distribution of *well-being*.

In this light, consider once again the example from the start of this section: in the first outcome, ten people have 100 units of well-being each; in the second, one person has 550 units, while the other nine have only 50 units each. In thinking about this case, it is natural to bring to bear our intuitions from real-life cases of distribution—treating the example as though it concerned the distribution of, say, units of money. Then the first outcome would indeed be better, since the total amount of well-being would be higher. But if we keep in mind that what the example actually involves is the distribution of well-being, and not the distribution of resources, we will see that these initial intuitions are misleading: by hypothesis the total amount of well-being is the *same* in both outcomes; there is, therefore, no reason to prefer the first outcome, no reason to prefer the outcome where well-being is distributed equally. Or so argues the welfarist.

But not everyone will be convinced by this argument. No doubt many will indeed find that when they reexamine their convictions, keeping in mind the potential confusions just noted by the welfarist, the commitment to egalitarianism has disappeared. But others may find that even when all of the welfarist's points are kept firmly in mind, they are still drawn to the thought that one outcome can be better than another by virtue of the fact that it contains a more equal distribution of well-being. They will insist that it is the welfarist who is mistaken in claiming that, so long as the total amount of well-being is the same, it is a matter of indifference from the moral point of view whether one person has a higher level of well-being than another. In short, they may remain convinced that equality matters—not only instrumentally, but intrinsically.

Of course, any committed egalitarian will need to say just *how much* equality matters. In our original example, both outcomes have the same total amount of well-being; they are “tied” from the perspective of the total view, and the egalitarian makes the (relatively modest) claim that the fact that one outcome has more equality than the other helps to break the tie. But sometimes an outcome with greater equality will have a *smaller* total amount of well-being. What should we say then? Is the importance of equality limited to breaking ties? Or can an increase in equality be valuable enough to outweigh a possible loss in terms of the total amount of well-being—so that on balance an outcome with greater equality but a smaller