

Natural Law

The Theory and Its Attractions

You are an animal.

I'm not trying to insult you. Just stating a fact. I am an animal, too. And so is everyone else we know. The basic needs of animals—food, water, security, companionship, freedom from pain—are the basic needs of human beings. All humans, like every other animal, share the same fundamental plight: certain one day to die, and vulnerable to harm in the meantime. Perhaps the key to morality lies in understanding our place in the natural order of things. Many have thought so.

In trying to discover what makes for a good human life, we might take a cue from the rest of the animal kingdom and ask about why their lives go well, when they do. It seems that there is a common answer. Animals live good lives when their nature is fulfilled, and bad lives when it isn't. A race horse, by nature, is built for speed. English pointers are meant to aid in the hunt. Chameleons naturally blend in with their background. When fillies break a leg, or chameleons cannot camouflage themselves, their lives go poorly. A good pointer will be able to track and give chase; a bad one will sit lazily and ignore nearby prey.

We can extend this sort of thinking even beyond animals. Impurities in a diamond render it flawed. A lake choked with algae is a poor lake. A torn aorta is a bad aorta. And so on.

In each of these cases, nature is dictating the terms of appraisal. The things *in* nature *have* a nature. Such things are bad when they are

unnatural, and good to the extent that they fulfill their nature. Perhaps we can say the same thing about human beings.

That is the guiding thought of the **natural law theory**. By its lights, good human beings are those who fulfill their true nature; bad human beings are those who don't. The moral law is the natural law—the law that requires us to act in accordance with our nature. (As we'll see, this is a different kind of natural law from the one that physicists use to describe the workings of molecules or galaxies.) At its most basic, natural law theory tells us that *actions are right just because they are natural, and wrong just because they are unnatural. And people are good to the extent that they fulfill their true nature, bad insofar as they flout it.*

The natural law theory promises to solve some very serious problems in ethics. Four of these are especially important.

1. *Natural law theory promises to explain how morality could possibly be objective, that is, how its standards depend on something other than human opinion.*

According to this theory, human nature can serve as the objective standard of morality. We do right when our acts express human nature, and do wrong when they violate it. Since individuals and entire societies can be mistaken about what our true nature is, they can be badly off target about what morality asks of us.

Although many natural law theorists are theists, who claim that our nature was given to us by God, that is not an essential element of the theory. What is crucial is that human nature is meant to serve as the ultimate moral standard. If this theory is correct, then so long as there is such a thing as human nature, there is an objective source of morality.

2. *Natural law theory easily explains why morality is specially suited for human beings, and not for anything else in the natural world.*

Almost everyone agrees that a distinctive human feature is our sophisticated reasoning abilities. A few other animals may be able to reason in basic ways, but no species on earth can approach our ability to assess various ways of life, critically analyze the merits of actions and policies, and then govern our behavior on the basis of our reflections. This capacity for rational thought also seems to be the cornerstone of morality. Moral agents—those who bear responsibility for their actions, and who are fit for praise or blame—are those who can control their behavior through reasoning. That's why we don't hold animals (or trees or automobiles)

morally responsible for the harms they sometimes cause. Only human beings have the sort of nature that enables them to be moral agents. Natural law theory can thus explain why moral duties apply only to human beings (or, if we ever discover them, to other life forms who share our rational powers).

3. *Natural law theory has a clear account of the origins of morality.*

The theory tells us that morality is only as old as humanity itself, that morality dates to the earliest days of humankind. But that isn't because morality depends on human opinion, as so many people believe. Rather, it is because morality depends on human nature. No humans, no morality. That's why there are no such things as eternal moral laws.

4. *Natural law theory may solve one of the hardest problems in ethics: how to gain moral knowledge.*

There are many skeptical arguments that try to undermine hopes for moral wisdom. Here is a perennial favorite, a variation on an argument developed by the brilliant Scottish philosopher David Hume (1711–1776).¹ Let's call this *Hume's Argument*, in his honor:

1. We can know only two sorts of claims: conceptual truths or empirical truths.
2. Moral claims are neither conceptual truths nor empirical truths.
3. Therefore, we can have no moral knowledge.

A **conceptual truth** is one that can be known just by understanding it. Here are some conceptual truths: No sphere is a cube; all integers are even or odd; bachelors are unmarried males; if A is taller than B, and B is taller than C, then A is taller than C. You can close your eyes to the world, just think about these claims and know that they are true.

Empirical truths are not like this. They are known only by relying on evidence from our five senses. Here are some empirical truths: I live in a house that was built in 1915; it was raining in London on June 25, 2007; the Pacific Ocean is larger than the Atlantic; David Hume never married.

Suppose we agree with the plausible claim that all knowledge is of either conceptual or empirical truths. If that is right, and if moral claims are neither, then moral knowledge is impossible.

Why aren't moral claims conceptual truths? Because for any moral claim, we can completely understand it and still wonder whether it is true. That doesn't happen when we ask whether bachelors are unmarried or

whether spheres are cubes. Anyone who really understands those questions already knows the answer.

Why aren't moral claims empirical truths? Because we can't see, touch, smell, taste, or hear them. Nor are they needed to explain anything that our senses inform us about. No amount of scientific probing into the world will reveal any moral features in it. If you witness a murder, or a broken promise, you will notice many things. But you can't see its *wrongness*.

Hume had a supporting reason for thinking that moral knowledge could not be empirical. When we describe the world, we talk about what *is* the case. But morality speaks of what *ought to be* the case. How can we get from descriptions to prescriptions? How does knowing how the world actually works enable us to learn how it ought to work? Hume thought that there was no answer to this question. If he is right, there is a gap between what is and what ought to be, a gap that can never be crossed.

Suppose an act is correctly described in the following way: as an intentional killing, one for which the killer felt no remorse, one that was extremely painful and prolonged, and inspired tremendous fear in its victim. If Hume is right, then no matter the number of descriptions we pile on, logic will never dictate any particular moral evaluation of such an action.

Of course we *think* that such a killing is immoral. But the claim that premeditated killing is immoral cannot be established by empirical evidence. Nor can conceptual truths establish it. Since all evidence must take one form or the other, it follows that we can never be justified in our moral views.

Natural law theory claims to be able to solve Hume's challenge. According to the theory, moral knowledge requires two things: we must know what our human nature is, and know whether various actions fulfill it. Natural lawyers (i.e., those who support the natural law theory) think that both kinds of knowledge are empirical. Human beings are part of the animal kingdom. We learn the true nature of other animals by careful scientific study. And the same holds for human beings. Discovering the essence of human nature is a scientific enterprise. Armed with this empirical knowledge, we can then look carefully at individuals to see whether their actions line up with human nature. This careful examination is empirical, too.

Suppose, for instance, that we undertake a vast study of human infants, across many different cultures, and discover that they are gentle and non-violent. Many have thought that this sort of empirical evidence clinches the case for thinking that these traits are part of human nature. If we then

see people acting aggressively and violently, we have all the empirical evidence we need to convict them of immorality, since they are undermining their true nature.

So, in the natural law view, gaining moral knowledge need not be mysterious. Armed solely with descriptions of a person's behavior, and knowledge of our human nature, we *can* determine whether actions are moral, by seeing whether they fulfill our nature. Just as we can tell whether a cheetah is a good member of its species by observing its speed, its talent for hunting, and so on, we can determine whether human beings are good or bad just by taking careful note of their actions and intentions. All we need is to conduct a scientific study of human nature, and then observe whether various actions fulfill that nature. This will amount to a good deal of empirical, descriptive knowledge, from which we can derive a moral conclusion. If natural law theorists are right, you really can derive an *ought* from an *is*.

Two Conceptions of Human Nature

We often voice our approval of actions by declaring them to be perfectly natural, or excuse someone's harmful conduct by saying that it was the natural thing to do under the circumstances. We also condemn certain actions as unnatural, or say of an especially awful act that it was a crime against nature. This all makes excellent sense, on the assumption that natural law theory is true.

The central claim of the theory is that the moral is the natural: The ideal for human beings is to fulfill their nature. Much depends, then, on what our nature really is.

In general terms, human nature is what makes us humans. It is the set of features that is essential to being human, so that if we were to lose these features, we would also lose our humanity. Natural law theorists are committed to the idea that there is a human essence, a set of traits that define us as human beings, and that mark us off as distinct from anything else in the world.

What is the nature of human nature? Here are two familiar—and problematic—answers.

Human Nature Is What Is Innately Human

Innate traits are ones we have from birth. They are natural in the sense of being inborn, natural as opposed to being learned, or acquired from

parents and society. In this line of thinking, our true nature is the one we are born with; traits we acquire through socialization are artificial, and stain the purity of our earliest days. In principle, we can use scientific methods to discover what is innately human, and so solve Hume's challenge to gaining moral knowledge.

If Jean-Jacques Rousseau (1712–1778) was right, we are innately angelic. Before society corrupts us, our noble nature shines through. We are by nature pleasant, cooperative, and considerate of others. If our nature holds the key to morality, then morality is largely as we think it is. It requires us to be kind, cooperative, and considerate, and forbids us from being the opposite.

This nicely supports traditional moral beliefs. But what if Thomas Hobbes (1588–1679) had it right? He thought that we are naturally selfish, competitive, and distrustful. We are born that way and, for the most part, stay that way. If the natural is the innate, and if we are required to act on our true nature, then the Hobbesian view is going to force us to abandon much of our conventional ethical beliefs.

The view that the natural is what is innate is widely held. This is what explains why so many people think that studies of animal behavior, or those focused on human infants, will unlock the key to human nature. The thought is that society is bound to change our natural state, and so we gain the deepest insight into human nature by discovering what we are like before society exerts its influences.

Yet if natural law theory is correct, and if the natural is the very same thing as the innate, then we need to solve the nature/nurture debate before we can know what is right and wrong. And that seems mistaken. We are *very* confident that morality is not a counsel of selfishness, mistrust, and competition, even if we are uncertain about whether such traits are innate. We can be very sure that killing people because of their skin color is immoral, even if we aren't sure whether we have an innate tendency to harm people who don't look like us.

This raises a general point: *The ultimate origins of our impulses are irrelevant to the morality of our actions.* Rape and robbery are immoral, no matter whether the impulse to commit these crimes is innate or acquired. Cheerfully comforting the sick is a good thing, even if we weren't born with a desire to offer such help. Since the morality of our actions and character traits does not depend on whether they are innate or acquired, natural law theorists must look elsewhere for an understanding of human nature.

Human Nature Is What All Humans Have in Common

Many people think that our nature is fixed by those things that we all have in common. These universal human traits would make up the essence of humanity. Such a view lets us scientifically determine our human nature. The data wouldn't always be easy to come by. But with a lot of effort, we could reveal the essence of human nature just by observing the features that every person shares.

There are two problems with such a view. First, there may be no universal human traits. And second, even if there are, they may not provide good moral guidance.

It may seem silly to deny that there are any universal human traits. Doesn't everyone want to have enough food and water to remain alive? Don't all adults have a sex drive? Aren't we all capable, to one degree or another, of complex thinking about our future? Yet some people want to die, not to live; others are indifferent to the attractions of sex; still others are so mentally impaired as to be unable to think at all about their future. For just about any trait (perhaps every trait) that is said to be part of human nature, we can find exceptions that undermine the rule.

Natural law theorists have a reply to this, which is best appreciated by considering an example. Return to the case of nonhuman animals, and think about their nature. For instance, it is part of a buck's nature to be alert to predators, to have four legs, to grow antlers, and to be fawn-colored. Still, there are bucks with only three legs. A few fail to grow antlers; others are deaf to predators; still others are albinos. We might say of such specimens that they aren't really bucks, not fully bucks, or not all that bucks should be.

If that sounds right, then we might adopt the following strategy. Perhaps human nature, like that of nonhuman animals, is determined not by what *every* member of the species shares, but only by what *most* members share. Bucks can have a nature, even if some bucks fail to perfectly live up to it. The same goes for human beings.

But this strategy won't work. There is the difficult problem of setting a threshold. Just how many humans need to have a trait before it qualifies as part of human nature? But leave that aside. The real problem is this: the fact that most humans have a certain trait is morally irrelevant.

Suppose, for instance, that most of us are selfish and vindictive. In this line of thinking, being selfish and vindictive would then be part of human nature. That would make such behavior morally right, on the natural law

view. But no natural lawyer accepts that. And they are right to reject such thinking.

Even if everyone, or most of us, were cruel and malicious, that would not make cruelty and malice morally good. Even if people were ordinarily, usually, or typically vindictive and petty, these traits would still be vices, not virtues. The fact that many, most, or all people behave a certain way, or have certain character traits, is not enough to show that such behaviors and traits are morally good. The line from *is* to *ought* cannot be crossed so easily.

Natural Purposes

If human nature is not a matter of the (innate) traits that all or most of us have, then what is it? The answer given by most natural law theorists is this: Human nature is what we are designed to be and to do. It is some function of ours, some purpose that we are meant to serve, some end that we were designed for.

It may seem that this conception of human nature places us squarely outside the realm of science, and in the domain of religion. For how could science tell us what our purpose is? Doesn't talk of our being designed for something imply the existence of an Intelligent Designer? In fact, many natural lawyers have made these assumptions, and have developed their views within the context of one religious tradition or another. But other natural law theories have been thoroughly secular. Such ethical views operate on the assumption that nature alone, unaided by anything supernatural, has endowed us with a purpose.

It makes sense to say that a Maserati has a certain nature: it is built for speed. It is perfectly plausible to say that it is in the nature of a Godiva truffle to be delicious, that it is meant to be tasty. When we speak like this, we mean that there is a person who creates such things with the intention that they serve a particular purpose. So, if it is part of human nature to serve some end, then this seems to imply that there is an intelligent designer who endowed us with this purpose.

And not just any intelligent designer. After all, Hitler's brilliant scientists designed the V-2 rocket in order to destroy English cities and kill thousands of civilians. If someone had discovered the rocket supply and destroyed it, the person would have undermined the purpose for which the rockets were created. But such "unnatural" action would hardly be immoral.

No, the intelligent designer in question must be God. In this line of thinking, God assigned us a specific purpose when He created us, and since God is perfectly wise and good, we do wrong when we act contrary to His designs. Frustrating God's purpose is immoral. That is why acting in a way that fulfills our nature is so morally important.

There is a lot to say about such a view, but most of it has already been said in the previous chapter. In the present account, we must act naturally because that is the way we respect God's plans for us, which are at the heart of morality. Though this isn't quite the same thing as making God's *commands* the basis of morality, it is close enough to have inherited most of the strengths and weaknesses of the divine command theory. Rather than revisit that topic, let's consider a secular interpretation of natural purposes.

The challenge is to make sense of the idea that we have been designed to serve some purpose, without having to invoke an intelligent designer. Strictly speaking, of course, nature has no designs for us. Nature is not an intelligent being with intentions and plans. Still, it *can* make sense to speak of something's natural function or purpose. We can cite the mechanisms of evolution and natural selection, rather than God, as the source of our natural purposes.

For instance, nature designed our brains to enable us to think, our liver to detoxify our blood, and our pancreas to regulate glucose levels. We can say what mitochondria are for, what the heart and kidneys are meant to do. In each case, there is a purpose that these organs serve, even if no one assigned them this purpose. But that sort of talk doesn't easily translate to human lives. What is a human being *for*? Does the question even make sense?

To answer this question, we need to understand the idea of a natural purpose. Two basic secular accounts might offer some insight. Call the first account the *Efficiency Model*, and the second, the *Fitness Model*.

Consider the Efficiency Model. Sticking with the example of a heart, we can say that pumping blood is its natural purpose, because nothing pumps blood as well as a heart. Hearts have a certain structure that enables them to pump blood more efficiently than anything else in the body. That is why the purpose of a heart is to pump blood.

Human beings can have a function or a purpose, then, if we are more efficient than anything else when it comes to certain tasks. Well, we are. But there are so many of them. We are, for instance, better than anything else at designing puzzles and writing essays. But on this model, natural law

theory cannot be correct, given its claim that unnatural action is immoral. For that would mean that we act immorally whenever we are inept at puzzle design or essay writing. We are also far better at building weapons than any other animal, and far more talented at using instruments of torture. But if acting naturally is always morally acceptable, then these actions, if they really are among our natural purposes, are beyond reproach. Something has gone wrong.

If the Efficiency Model is correct—if human nature is given by our natural purposes, and these purposes are whatever we are best able to accomplish—then natural law theory must fail. There are too many such purposes, and many have nothing moral about them. Perhaps the Fitness Model will do better.

On this account, our organs have the purposes they do because it is extremely *adaptive* for them to serve these roles. The natural purpose of the heart, brain, liver and lungs is to do what enhances **fitness**: roughly, our success at survival and reproduction. We are able to survive, and pass on our genes to our offspring, only because these organs function as well as they do. Nature has designed hearts and kidneys and brains (etc.) to improve our chances of survival. This is their natural purpose; it is ours, too. We are meant to survive, and to transmit our genes to the next generation. That is what a human life is *for*. In a godless world, that is all the purpose our lives can have.

Since our natural purposes are survival and procreation, we can see why so many natural law theorists have thought suicide immoral, and have condemned birth control and homosexual activity. We also have a ready explanation of why courage, endurance, and fortitude are true virtues—those who possess them are (in the relevant sense) fitter than those who don't.

Suppose that the natural law theory is true. And suppose that we fulfill our human nature just when we fulfill our natural purposes. Two things follow:

1. Acting naturally—fulfilling our natural purpose—is always moral.
2. Acting unnaturally—frustrating our natural purposes—is always immoral.

But if the Fitness Model is correct, then both claims are false.

Claim 1 is false. To see this, recall that natural actions are those in which we use our mind and body to satisfy the purposes they were designed for. In the Fitness Model, these purposes are survival and reproduction.

So natural actions are those that increase the chances of our survival and reproduction. But men can increase the chances of passing on their genes by raping as many women as they can. That is about as immoral as anything I can think of. And survival? Consider the words of Primo Levi, an Auschwitz prisoner: "the worst—that is, the fittest—survived. The best all died."²² Sometimes those best schooled in treachery, intimidation, and violence are the ones likeliest to live another day. If we understand natural purposes as the Fitness Model advises, then claim 1 is false.

Claim 2 is also false. Not every act that frustrates a natural purpose is immoral. Nature has engineered our ears to be capable of hearing—the better to detect predators, to listen to the advice of our allies, to hear the threats posed by our attackers. But there is nothing immoral about wearing a set of headphones that block out noise. We have eyes so that we can see. But there is nothing wrong with crossing your eyes to make a joke, or closing them to shut out an unwanted sight.

It is worth noting that these examples can be successful even if it is God, and not nature alone, that has endowed us with these various purposes. Suppose that God made eyes to see, ears to hear. Still, isn't it morally acceptable to put on blindfolds, or wear headphones? Despite being "unnatural," these actions are perfectly acceptable.

What this shows is that the Fitness Model is as vulnerable as the Efficiency Model. Neither gives us a solid understanding of what human nature is. Until we are given a better method for determining our nature, the natural law theory is in trouble.

The weakness of the various understandings of human nature allows us to see why a classic moral argument fails. That argument—call it the *Natural Law Argument*—goes like this:

1. If an act is unnatural, then it is immoral.
2. Suicide, contraception, and homosexual activity are unnatural.
3. Therefore suicide, contraception, and homosexual activity are immoral.

The first premise is false on all of the interpretations we have so far considered. Whether unnatural actions spring from acquired traits, rather than innate ones; whether they are rare or unusual, rather than typical or even universal; whether they frustrate nature's purposes rather than conform to them; still, such actions can be morally acceptable.

This does not prove that suicide, contraception, or homosexual activity are morally okay. What it shows, however, is that this popular argument

designed to reveal their immorality is highly suspect, and will certainly fail unless we have a better understanding of human nature to rely on.

The Argument from Humanity

This may be a good time to consider another famous argument of the natural law tradition: the *Argument from Humanity*. This is perhaps the most widely voiced anti-abortion argument in public debates on the subject. The argument is straightforward:

1. It is always wrong to deliberately kill an innocent human being.
2. A fetus is an innocent human being.
3. Therefore, it is always wrong to deliberately kill a fetus.

A great deal could be said about that first premise, but let's leave it aside for the moment, and focus on the second. It seems clearly true to a great many people. And it seems just as clearly false to many others. What is going on here?

The explanation, I think, is that the term *human being* is **ambiguous**—it has more than one meaning. There are at least two senses of the term used in these debates, and partisans in each camp tend to use their preferred meaning. The result is that a lot of the discussions about abortion end up going nowhere.

For our purposes, the important question is whether a fetus is an innocent human being. And that depends on how we define humanity. What is the essence of humanity? It is tempting, of course, to look to science for an answer. It tells us that being a member of the species *Homo sapiens* is the essence of humanity. On this biological account, premise 2 is clearly true, since fetuses of our species are certainly innocent of any wrongdoing.

But if we give a purely scientific definition of humanity, then premise 1 **begs the question** against pro-choice opponents. In other words, premise 1 does not provide an independent reason for rejecting the pro-choice position. Those who advance premise 1 without any supporting argument for it are preaching to the choir, since only those who already oppose abortion will accept this first premise. If humanity is defined in purely biological terms, then the first premise needs a lot of defense—indeed, as much defense as the argument's conclusion.

Alternatively, we might think of humanity not as a biological category, but rather as a moral one. On such a view, to be human is to have a certain moral status. It is to be entitled to an extensive set of moral rights, including the right to life. On this reading, premise 1 may well be true, though

some critics will argue that even this moral rule has some exceptions. (We consider the matter in some detail in the chapters to come.) Still, even if premise 1 turns out to be true, the second premise, the one that grants fetuses a wide range of basic rights, just as clearly begs the question. It needs as much defense as the conclusion it is meant to support.

So the term *humanity* (like the term *natural*) is ambiguous. This isn't any kind of problem, so long as we are very clear about which meaning we are relying on. But once we *are* clear about its different meanings, and make sure that the same meaning is being used in both premises, we can see that the argument is bound to beg the question. What this shows is that the argument cannot stand alone. Depending on which meaning of *human* we use, opponents of abortion will have to provide a strong supplemental argument to defend premise 1, premise 2, or both.

There are lots of cases like this. Here is another example, from a newspaper article I read just this morning:

1. It is morally okay to play sports.
2. Dogfighting is a sport.
3. Therefore, dogfighting is morally okay.

This argument was given by a friend of Michael Vick, the star quarterback who pled guilty to criminal charges related to his role in running dogfights. Before you had a look at this one, you might not have realized that people have different definitions of sport. But they do. Some people think of a sport as any skill-based athletic competition. Others think of it as any recreational activity. Still others consider sports to have essentially character-building and morally uplifting aspects, so that a gladiator contest, for instance, or professional wrestling, doesn't qualify as a true sport.

None of these definitions is uniquely correct. The term *sport* is also ambiguous. And that's okay. It's a terribly impoverished language that can't make room for ambiguity. But when evaluating a line of argument that contains an ambiguous term, we must first settle on the definition we will use, and then stick with it. Only then can we test each premise to see whether it turns out to be true.

When we do that with the dogfighting argument, we can see that those who think that sports are essentially character-building will accept the first premise, but likely reject the second. And those whose definitions make no reference to morality may accept the second premise, but then, when they contemplate so-called "blood sports," cast a suspicious eye on

the first. The same pattern that we saw in the Argument from Humanity repeats itself here, and in so many other contexts.

There is a basic philosophical point here: You can't solve complex moral problems with definitions alone. You can't solve the abortion debate just by defining a human being in terms of species membership. Nor can you solve it if you define a human being as a person possessed of basic moral rights. We can define *humanity* in any way we like—by reference to species, genetic code, moral rights, or to powers of reasoning, self-awareness, linguistic ability, or in any of a dozen other ways. But no matter what definition we come up with, there is no shortcut through a lot of further complicated moral argumentation. Definitions alone will never spare us the hard philosophical labor it takes to solve complex moral problems.

There is no better contemporary illustration of this point than the *Marriage Argument*, given by many natural law theorists and other opponents of same-sex marriage:

1. Marriage is defined as a relation between a man and a woman.
2. Homosexual relations are between men and men, or women and women.
3. Therefore, homosexual relations can never qualify as a marriage.

The search for a definition is the search for an essence. Defining bachelors as unmarried adult males is supplying the essence, the central nature, of being a bachelor. Is marriage essentially a relation between those of the opposite sex, as defenders of the argument claim? Or is its essence the expression of mutual love, honor, and commitment, as supporters of gay marriage assert? Or something else entirely?

I won't try to settle this debate here. But the point raised earlier applies to this argument as well. The real question behind today's debates about gay marriage is whether the state is morally required to extend the same legal rights to same-sex couples that it now grants to heterosexuals. *That* question cannot be settled by setting forth a definition, especially one that will appeal only to those on one side of the debate. It can be settled only by examining the ultimate point of marriage, the legitimate role of the state in supporting various practices, the morality of homosexuality, and other difficult matters.

Definitions are tools for thinking clearly. They clarify the subject matter and tell everyone precisely what the focus of discussion is. But having a subject firmly in mind is one thing; determining its moral status is quite another. Definitions alone will never solve a moral problem.

While this does not undermine the natural law theory, it must make us very careful of how we employ it. Even if we agree on a definition of human nature, and so on a specific understanding of our essential traits, there is a great deal to do—indeed, there is almost everything yet to do—before we can draw important moral lessons about how to live our lives.

Conclusion

The deep appeal of the natural law theory is its promise to base morality on something clear and unmysterious: nature and its workings. Moral laws, on this account, are just natural laws, though ones that regulate human beings, rather than planets, molecules, or gravitational forces. But as we have seen, it is terribly difficult to try to read off recommendations for how we ought to act from descriptions of how nature actually operates.

And that shouldn't be too surprising. Natural laws describe and predict how things will behave. They summarize the actual behavior of things, and, unless they are statistical laws (of the sort that assign a probability to outcomes, rather than a certainty), they cannot be broken.

Moral laws are different in every respect. They can be broken, and often are. They are not meant to describe how we actually behave, but rather to serve as ideals that we ought to aim for. Nor are they designed to predict our actions, since we so often fall short of meeting the standards they set.

Nature can define the limit of our possibilities. Our nature does not allow us to leap tall buildings in a single bound, or to hold our breath for hours at a time. On the assumption that morality does not demand the impossible of us, nature can, in this way, set the outer bounds of what morality can require. But it can do no more. It cannot, in particular, tell us what we *are* required to do. Nor can it tell us what we are forbidden from trying to achieve. Nature has, at best, only a limited role to play in moral theory.

Notes

1. The original argument appears in his *Treatise on Human Nature* (1739), Book III, part I.
2. Primo Levi, *The Drowned and the Saved* (New York: Knopf, 1986), p. 82.

Psychological Egoism

Egoism and Altruism

Early on in his masterpiece, the *Republic*, Plato recounts a story that was well-known to his readers. This is the tale of Gyges the shepherd. Trying to catch up with his wandering flock, Gyges enters a cave. On its dirt floor he finds a golden ring. He puts it on and discovers that it makes him invisible. After testing it a few times to make sure of its powers, Gyges goes to town, kills the king, captures the throne, and weds the queen. He is able to satisfy his every want. And does.

The lesson many have taken from this story is that we are always looking out for number one. If we had the power to do as we liked, free of any worries about punishment, we would always seek out our own best interests, no matter the harms we caused. The only thing that keeps us in check here in the real world, where no such ring exists, is the worry about how we'd be made to suffer if we did the things we most wanted to do.

In civilized society, no one is all-powerful. We each offer help to other people, some more, some less. But this doesn't change our basic motivation. We see that in a world of limited power, we do best for ourselves by sometimes helping others. We keep our word so that we gain the trust of those who can help us. We don't kill our enemies, for fear that we will be the next target. But no matter our position in the world, we all want the very same thing: to make ourselves as well off as we can be.

This is the view of **psychological egoism**, which tells us that there is only one thing that motivates human beings: self-interest. If this theory is

true, then **altruism**—the direct desire to benefit others for their own sake, without any ulterior motive—does not exist.

And yet I have just visited a website, <http://carnegiehero.org>, which records the stories of everyday people who have undertaken life-saving rescues at extreme personal risk. There are many reports of people jumping into freezing waters or overturned and blazing automobiles in order to save complete strangers. A number of the rescuers died in their efforts.

Why would people do such things? There seem to be many reasons. Some tried only to help a victim in need. Some felt called by duty to offer help. Others may have had mixed motives, of compassion and care on the one hand, and perhaps a small desire for praise and popularity on the other.

If psychological egoism is true, then these explanations are all wrong. None of these people was a true hero, since every one of them was driven by self-interest to save the lives of others. It is important to see how strong a claim this is. Psychological egoists are not saying that people usually act out of self-interest, or that altruism is rarer than it seems. Rather, their claim is that actions are *never* done from altruistic motives. This isn't just a fluke, either. It's not as if we are capable of being altruistic, but never decide to go that route. Rather, the psychological egoist explains the absence of altruism by claiming that altruism is *impossible* for us human beings. We can't fly. We can't live in five-thousand-degree temperatures. And we can't be devoted to anyone other than ourselves.

True, people often do help one another. But egoists will explain this in self-interested terms. Suppose, for instance, my boss asks me at the last minute to baby-sit his bratty children. I am miserable, but smile anyway and quickly agree. Though I benefit my boss, the only reason I do so is because I expect to get something from it. If I didn't—if I knew, for instance, that he was soon retiring, and was powerless to advance my career or get me a raise—then I wouldn't agree to help. Psychological egoists are well aware that people help one another in many ways. But that doesn't show that they do so from altruistic motives.

As its name implies, psychological egoism is a psychological view, rather than an ethical theory. It aims to *describe* the facts and limits of human motivation, rather than *prescribe* the standards that we ought to live up to. If it is true, every single action—mine and yours and everyone else's—is done from the hope of personal gain.

This view strikes many as revealing a deep truth about ourselves. And it strikes others as the height of cynicism. Those who prefer a more

flattering view of humanity, one that allows us our altruistic motives, will want to debunk the theory. It isn't easy.

We can't refute this kind of egoism just by pointing to examples, such as smoking, in which people know that they are jeopardizing their health or their life. That is because those who smoke presumably gain some satisfaction from doing so. They are moved by a desire for pleasure, even if it lasts only the three or four minutes that their cigarette does.

Nor can we refute egoism by pointing to examples of people accepting extreme pain—say, the sort that comes from a root canal. When people sit still in a dentist's chair, they do so with their future in mind. They want to prevent something really bad from happening down the road—gum disease, rotting teeth, etc. It's hard to imagine why people would accept such pain if they thought there was nothing to be gained by it.

So the egoist's claim must be understood as follows: *All human actions are aimed at avoiding some personal loss or gaining some personal benefit (or both), either in the short run or in the long term (or both).*

Many actions fail to achieve any of these aims. This doesn't undermine egoism, though, since the theory is not about the results of our actions but about their motivations. An action can be done with the thought that it will gain you millions, even if it leaves you in debtor's prison. Such a case is no threat to psychological egoism.

It is true that certain kinds of behavior, such as sneezing or snoring, are not done from self-interest. Egoists can handle this sort of case as well. The egoist thesis is one about actions that result from a person's intention to do something. Suppose that an epileptic suffers a seizure. Her arm shoots out and hits someone. The egoist needn't say that she was somehow trying to benefit herself. She wasn't intending to do anything at all. In order to refute psychological egoism, you must provide examples in which a person intends to do something, even though she wasn't trying to benefit herself.

There are three popular arguments designed to show that no such example can be provided. But before we consider those arguments, we need to ask a preliminary question. Since psychological egoism is not an ethical theory, what bearing could it have on our ethical concerns?

As it happens, quite a bit. The *Implications of Egoism Argument* spells this out pretty clearly:

1. If psychological egoism is true, then we can't be altruistic.
2. If we can't be altruistic, then it can't be our duty to be altruistic.

3. Therefore, if psychological egoism is true, then it can't be our duty to be altruistic.
4. Psychological egoism is true.
5. Therefore, it can't be our duty to be altruistic.

Premise 1 is true by definition. No matter whether you like or hate psychological egoism, you should accept this premise. Premise 2 is also very plausible. If we can't be altruistic, then it can't be our duty to be altruistic. Why? Because we are not required to do the impossible—morality might be pretty demanding at times, but it can't be *that* demanding. 3 follows logically from 1 and 2, so if they are true, as they certainly seem to be, then 3 must be true as well.

That leaves only premise 4, which asserts the truth of psychological egoism. Suppose that it is true, and that altruism is a myth. Then we have no duty to be compassionate, benevolent, considerate, kind, or generous. We would have to radically change our moral ideals, ridding them of altruistic elements. The resulting morality would be largely unrecognizable. Most of what we take for granted about the ethical life would turn out to be mistaken.

It may sound overdramatic to put it this way, but the fate of morality as we know it depends on whether psychological egoism is true. We must give up our ordinary understanding of what morality demands if psychological egoism turns out to be correct. Let's see whether it is.

The Argument from Our Strongest Desires

There are three central arguments that seek to establish the truth of psychological egoism. The first of them is the *Argument from Our Strongest Desires*. The argument begins by claiming that every action you perform is based on your strongest desire. But if your strongest desires are what's moving you, then you are pursuing self-interest. So, whenever you act, you are pursuing self-interest. That is just what the psychological egoist claims:

1. Whenever you do something, you are motivated by your strongest desire.
2. Whenever you are motivated by your strongest desire, you are pursuing your self-interest.
3. Therefore, whenever you do something, you are pursuing your self-interest.

Consider the argument's initial claim, that our actions are always caused by our strongest desires. Apparent counterexamples spring to mind: visiting a nagging relative, enduring a tedious commute, handing over your wallet to a gunman. Does anyone really want to do these things? Probably not. But that doesn't show that this first premise is false. For we can explain such actions by pointing to *other* things that people want—to sustain family harmony, to keep one's job, to stay alive.

The case of **strictly conscientious action** is more difficult for the egoist to handle. Such action occurs when you do what you think is required of you, even in the face of great temptation. Your desires all push one way, and duty calls in another. People sometimes follow the latter path.

Of course, we could say that even here, people are acting on their strongest desire. It's just that in cases where people knuckle down and resist temptation, their strongest desire is to do their duty.

This move could save premise 1. But that should be cold comfort to egoists. For if they relied on this move, then they would have to admit that premise 2 is false. They would have to say that our deepest desire is not always to benefit ourselves. Sometimes we most want to do our duty, even when it comes at a personal cost.

When we put it this way, it becomes clear that egoists must deny the existence of strictly conscientious action. Of course, people do help others, and people sometimes speak of themselves as acting just for the sake of duty. But egoists do not take such talk at face value. They insist that those who help others, despite what they might say, are really just looking out for themselves. People sometimes lie to others to impress them. And people sometimes lie to themselves, deceiving themselves with self-flattery.

These points are undeniable, but they fail to show that conscientious action never occurs. We can all think of apparent cases of such action, and even if some of the evidence is tainted by lies and self-deception, we have no reason (as yet) to think that *all* of it is suspect. Until we are given such a reason, the case of conscientious action presents a real problem for the egoist's argument.

There is another way to put this worry. Let us grant for the moment that premise 1 of the argument is true. So we always do what we most want to do. But *that doesn't yet show that our strongest desires are always for personal gain*. That is precisely what has to be shown.

Have a look at premise 2 again. It states that our strongest desires are always for our own personal gain. But those who believe in altruism will not accept that. Indeed, if we sign on to premise 1, then premise 2 begs

the question—it assumes the truth of the conclusion that it is meant to support. It is preaching to the converted. It is not a neutral thesis that can appeal to both fans and opponents of psychological egoism. Only those who already reject the existence of altruism will accept it. Everyone else will think that our strongest desires may sometimes be for something other than self-interest. We therefore cannot accept premise 2 without a lot more argument.

Premise 2 assumes that just because a desire is mine, it must have a certain object—me, and my self-interest. But *whose* desire it is, and *what* the desire is for—these seem to be completely separate issues. Since we are assuming now that premise 1 is true, we can say that whenever I act, I do so from my strongest desire. But that does not settle the question of what my desire is *for*. The egoist says that just because a desire is mine, then it has to be aimed at my self-interest. But why couldn't it be aimed at your welfare? Or the well-being of a friend, or my country, or a sports team, or even a stranger?

Indeed, some people seem to have malicious desires, aimed at hurting others, rather than at benefiting themselves. They may feel good about imposing such harm, but they may not. Some people know that they are doing wrong, know that they are going to regret it, and yet set out to harm other people anyway. When they do, we can grant that they are acting on their strongest desire. Yet it isn't a self-interested one—especially if, as sometimes happens, one deliberately harms oneself in order to harm someone else. (Think of the spurned boyfriend who sets out to kill himself, motivated by the thought of how painful this is going to be to his ex-lover.)

The egoist might respond: If you are doing what you really want, aren't you thereby self-interested? It is important to see that the answer may well be no. For all we know, some of us deeply want to help other people. When we manage to offer such help, we are doing what we really want to do. Yet what we really want to do is to benefit someone else, not ourselves.

Now, if people get what they really want, they may well be better off as a result. (But they might not: think of the anorexic or the drug addict. Think of the cases of disappointment discussed in chapter 4, p. 49) Yet the fact that a person gains from her action does not prove that her motives were egoistic. The person who really wants to help the homeless, and volunteers at a soup kitchen or shelter may certainly derive pleasure from her efforts. But this doesn't show that pleasure was her aim. Her aim may have

been to help those in need. And because her aim was achieved, she thereby received pleasure.

As a general matter, when you discover that your deepest desires have been satisfied, you often feel quite pleased. But that does not mean that your ultimate aim is to get such pleasure. That's what needs to be shown; we can't just assume it in trying to figure out whether our motives are always self-interested.

In short, even if this argument's first premise is true, its second premise begs the question. We don't yet have reason to doubt the possibility of altruistic motivation.

The Argument from Expected Benefit

So let us consider another argument, the *Argument from Expected Benefit*. This argument claims that people always expect their actions to leave them at least a bit better off. And if they do, then their constant aim is to gain from their actions. So whenever people act, they are trying to get some personal benefit. That is just what psychological egoism predicts:

1. Whenever you do something, you expect to be better off as a result.
2. If you expect to be better off as a result of your actions, then you are aiming to promote your self-interest.
3. Therefore, whenever you do something, you are aiming to promote your self-interest.

I have my doubts about this argument. Premise 1 seems to ignore the existence of pessimists. And even optimists sometimes expect to suffer for their actions. Consider a person who thinks she can get away with a convenient lie, but admits the truth anyway, knowing the misery that's in store for her as a result. Or imagine an employee late for an important appointment who increases his delay by helping a stranger cross a dangerous street. He doesn't anticipate any reward for his good deed, and knows that this delay is only going to stoke his boss's anger. Both cases seem to be counterexamples to the claim that our actions are always accompanied by an expectation of personal benefit.

These examples, and the many others like them that we could imagine, are bound to be controversial. Egoists will insist that people always *do* expect to gain from their actions, even if it sometimes appears that they don't. After all, appearances can be misleading. It may look as if we

sometimes expect only the worst. But deep down, we may always believe that our actions will make us better off.

I am not convinced, but for now, let's assume that this reply works and that premise 1 is secure. Even so, the second premise—the one that says that if you expect a benefit, then that is your aim—is very implausible.

The problem is that it looks like the egoist is begging the question again. Return to those who enjoy volunteer work. Such people may well expect to gain something from their activities. Volunteers often report feelings of deep satisfaction at their efforts. But this doesn't show that their motives are self-interested. A volunteer, with all sincerity, may claim that her aim is to help others, rather than to help herself. The enjoyment she experiences from her volunteer activities is nice, no question about it. But it is a side effect, a foreseen benefit, and not the motive of her actions.

Again, the egoist will insist that we can't be sure of her sincerity. And even if she is sincere, she may not realize the true motives that get her to the shelter every week. She may just be looking for the approval of her fellow volunteers, or for her efforts to be publicized in a way that will boost her career. We can't be sure that this isn't what's going on.

That is true: We can never be absolutely certain of our own motives, much less those of others. But that is not enough to show that the egoist is correct. Her claim—that whenever we expect some benefit, our aim is to get it—is controversial. It *seems* that there are counterexamples, though we can't be sure there really are. Given evidence that appears to undermine premise 2, the egoist needs to defend it.

And a defense is not far to seek. The best way to support this premise is by citing a more general principle that supports it. And this is the natural candidate:

(G) Whenever you expect your action to result in X, then your aim is to get X.

If G is true, then whenever I expect to get money, glory, or approval from my actions, that is my aim. Generally, if I expect some personal gain, then my goal is to obtain it. That is just what the argument's second premise says.

The problem is that G is false. Whenever I lecture to a large audience, I expect some people to fall asleep. Believe me, that is not my goal. If I ever had the chance to play against a professional tennis player, I'd expect to lose. But it wouldn't be my aim to do so. My goal would be to enjoy the

experience, and to learn a thing or two. If a student fails to prepare for an exam, she may expect to receive a poor grade. It hardly follows that she is trying for one.

So expecting something doesn't always mean aiming for it. And that poses a direct challenge to premise 2 of the Argument from Expected Benefit. It might be true, but we have seen no good reason to think so. Indeed, the best reason to think that 2 is true—principle G—is false. If we are to have any confidence at all in this argument, the egoist needs to furnish us with much better support for both of its premises.

The Argument from Avoiding Misery

Still, many people feel the force of the egoist's assertion: If an action promised only misery, then we wouldn't do it. Thus our actions are always prompted by a desire to avoid such misery. This is a self-interested motive. Therefore, all of our actions are prompted by self-interest.

This is a variant on the previous argument, but it has been very influential in convincing people to adopt psychological egoism, so it is worth spending a little time with it. We can reconstruct the *Argument from the Avoidance of Misery* as follows:

1. If we would never do an action that promised only personal misery, then all of our actions are done in an effort to avoid such misery.
2. We would never do an action that promised only personal misery.
3. Therefore, all of our actions are done in an effort to avoid personal misery—and that is a self-interested motivation.

Premise 2—the claim that we'd never act in ways that promised only misery—may well be false. Consider a gymnast who gives up a place on her country's Olympic team to allow a fellow athlete the coveted spot. Or a prisoner of conscience who stands by her principles, knowing that she is going to be tortured as a result. Or a killer who wants only to suffer, in order to repent for his crimes.

Egoists will try to point to benefits that each sacrifice is expected to bring. Perhaps they will always succeed. But even if they do, that wouldn't show that egoism is true.

That's because premise 1 is highly suspect. It is supported by this principle:

- (P) If I would never do an action that promised me only X, then I am always trying to avoid X.

The egoist says that we wouldn't do actions that promised only misery; therefore we are always trying to avoid misery. And avoiding misery is self-interested; therefore we are always motivated by self-interest.

But there are many counterexamples to P. I wouldn't do an action if it guaranteed me only a thousand bee stings, or permanent bankruptcy. That doesn't mean that I am always trying to avoid these things. I wouldn't do an action if the only possible outcome were a lifetime supply of hard-boiled eggs and marshmallows. (I hate these things.) That hardly means that my every action is aimed at avoiding them. Perhaps it's also true that I'd never do an action that promised me only certain death. But that doesn't mean that my main motive in reading philosophy or watching *Arrested Development* is to avoid dying. Thus even if we wouldn't do an action if it promised only misery, it just doesn't follow that our motivations are always to avoid such misery.

And so even if you don't like my earlier counterexamples to premise 2, the problems with premise 1 threaten the entire argument. If we are to accept the argument's conclusion, both premises must be credible. At this point, we have good reason to doubt each of them. Without further defense, we should withhold our support from the Argument from Avoiding Misery.

Two Egoistic Strategies

Things aren't looking that promising for egoists. But there are still two strategies that they can use to try to show that the evidence of altruism is overrated.

Appealing to the Guilty Conscience

Consider people who have taken great risks to oppose oppressive regimes. Many of these people claim that their conscience wouldn't let them do otherwise—had they taken the safe path, they wouldn't be able to live with themselves. In their eyes, to give in to evil is to tarnish oneself. Many people speak of the terrible guilt they'd feel if they did nothing to fight against injustice.

Egoists insist that even these people are wholly self-interested. They are opposing injustice in order to make sure that they can sleep well at night, that they can be free of crippling guilt. Having a clean conscience is a benefit. And so such people are acting from self-interested motives.

It is important to see why this sort of reasoning does not work. If a person is truly good, she will certainly be troubled at the thought of doing

wrong. But that does not prove that her actions are motivated by a desire for a guilt-free conscience. Indeed, if she did not care about others, then she wouldn't lose a wink of sleep at the thought of their misery. Those who suffer pangs of guilt from having harmed others, or having missed a chance to help them, are precisely those who care about other people.

There is a basic difference between those who are deeply upset over the damage they cause and those who aren't. It is reported that Eichmann slept like a baby, despite having engineered the deaths of hundreds of thousands in the concentration camps. And we all know others who are beside themselves at the very thought of having done a serious wrong. Egoists lump both kinds of people together and, in doing so, fail to appreciate that those who are subject to guilt must have altruistic concerns. If they didn't, then they wouldn't feel so bad when failing to help others in need.

Expanding the Realm of Self-Interest

Consider the mother who gives away the last of her food to save her only child. This seems like the essence of altruism. And yet the egoist might say that the mother is really looking out for herself, by trying to avoid a terrible personal loss. For she would be devastated at witnessing the death of her child. Further, since the mother cares deeply about her child, the mother's well-being rests in large part on how well her child is faring. By helping her child, the mother is thereby helping herself.

Much of what was just said is true. But this cannot be good news for the egoist, since the details of this little story imply that egoism is false. For most parents, their own well-being crucially depends on that of their children. And so, when parents tend to the needs of their children, they are usually helping themselves in the bargain. But this doesn't show that parents are motivated by self-interest when they offer such help. As we've seen, even if people expect to gain by helping others, that doesn't prove that their aim is to acquire such benefit. Further, if a parent suffers at the thought of her child's misery, then that is evidence of altruism, not egoism. Those who care only for themselves do not suffer when thinking of the misery of others.

The second egoist strategy is thus confused. When we present difficult cases for egoism—that of the soldier jumping on the grenade in the crowded foxhole, that of a mother giving her last bit of nourishment to her child—the egoist cannot sensibly reply that such people are motivated only by self-interest. For such people, what really matters is living honorably, or seeing their children flourish. Their concerns are directed at something

other than their own gain. That is not something that the egoist can make sense of.

Letting the Evidence Decide

People often benefit one another. Just think of the good that aid workers, nurses, teachers, and parents do. Sometimes their motives are clearly self-interested. But not always. In many cases, people report that their primary motivation is to help someone else, even when such action is known to come at a personal cost. As we have seen, the egoist must view all such reports as misleading, based on either self-deception or the intended deception of others.

This puts the egoist's thesis in jeopardy. We can see the danger by constructing a dilemma: Either we do or we don't allow the evidence to settle the question of whether altruism exists. There is trouble either way.

To appreciate the difficulties, consider this cautionary tale. As I write this, my hair, once dark black, now has a substantial amount of gray in it. Suppose a friend of mine starts teasing me about this. I tell him that I can't help it. "Yes, you can," he replies. "Grecian Formula will do the trick." "Oh, I doubt that," I say. "Those elves won't let anything stand in their way."

What elves? Why, the invisible ones. The ones who have turned my black hair gray. Each night they cover a strand or two with permanent gray paint, and over time, I've become more gray than black. And worse, there's nothing I can do about it, since I can't catch them and put a stop to the madness. After all, they are invisible. And very clever.

My friend, who thought I was joking, now becomes alarmed. "You know, don't you, that invisible elves don't exist. And hair color is a matter of genetics." He tracks down the latest data to make his point. Naturally, I dismiss this as the product of a conspiracy planted by the geneticists. No matter what sort of evidence my friend provides, I stick to my guns—it's the elves who are behind it all.

This unshakeable conviction is not the mark of a rational mind, but rather the crazy faith of someone seriously out of touch with reality. In refusing to allow for the possibility that I may be mistaken, I am being stubborn and dogmatic. If the evidence fails to support my beliefs, then I fault the evidence, rather than change my beliefs.

Some people are utterly convinced that the earth is flat; others, that space aliens built the Egyptian pyramids; yet others, that George Bush ordered the destruction of the World Trade Center. Many of these people

feel certain of their views. No amount of evidence is going to get them to change their mind.

But these beliefs, like mine in the invisible elves, are not rational. They are not supported by the evidence. And the worry is that a belief in psychological egoism is no better than the ones I've just described. If psychological egoists stubbornly deny all contrary evidence, and routinely dismiss the many cases of apparent altruism, then their view is similar to the invisible elf hypothesis. If I refuse in advance to allow the chance that my pet theory is mistaken, and refuse even to admit the possibility that evidence might cast doubt on it, then I am irrationally clinging to a prejudice.

For all I have said, psychological egoism may yet be true. And invisible elves may really exist. But it wouldn't be rational to think so.

So let us opt for the other half of the dilemma. We should keep an open mind and let the evidence decide whether altruism exists. But if we go this route, and subject psychological egoism to serious testing, then egoism fails the relevant tests.

Egoism is a theory about human motivation. And there are two sorts of evidence we can rely on when it comes to determining people's motives: testimony and behavior. People can tell us what moves them, and we can see for ourselves how they behave. Neither kind of evidence is perfect. We can misinterpret behavior. And, as we've already seen, people can be deceptive about their true motives. Still, testimony and behavior are the only sources of evidence we have.¹ And they both point strongly to the existence of altruism. To discredit all such testimony, and to reinterpret all such behavior, is to render egoism as unreasonable as my invisible elf hypothesis. If we are prepared to let evidence decide the matter, we must decide that psychological egoism is false.

Conclusion

The truth of psychological egoism would spell the defeat of morality as we know it. For if altruism is impossible, then morality cannot reasonably ask us to sacrifice self-interest for the sake of others. The central moral virtues of benevolence, kindness, and compassion would have no place in morality. If all we can do is look out for number one, then there is little point in demanding that we do otherwise.

So the stakes are high. And psychological egoism can seem the only clear-eyed, unsentimental view of human nature. We're all familiar with

cases in which we gave ourselves credit for being altruistic, only to realize, later on, that we were really just pursuing our self-interest.

But when we examine the arguments and evidence, there is reason to take a more charitable view of ourselves. Few of us are saints. But few of us are wholly self-absorbed, either. And this opens up the possibility that we may become better than we are, more altruistic and more generous, shifting our attention away from our own needs and onto the needs of others.

A plausible philosophical principle, first offered by Aristotle, is instructive in this context. The principle tells us to *follow the appearances*. It directs us to think that things are really as they seem, that appearances match reality—until we have excellent reason for doubt. It *seems* that there is plenty of altruistic motivation. People often speak of how much they care about others. And there are many cases in which people have actually offered help to others, strangers and loved ones alike. All of this evidence might be highly misleading. But the arguments that egoists provide, and the strategies they offer to handle counterexamples, are unconvincing. They fail to give us excellent reason to doubt the appearances of altruism. Until better arguments and strategies come along, we should trust the appearances and regard psychological egoism with great suspicion.

Notes

1. What about sophisticated fMRI or other brain-scanning techniques? They may someday provide evidence, but any such evidence really depends on the evidence of testimony and behavior. For how will scientists be able to tell whether a certain synapse firing indicates altruistic motivation? The only way to do that is to correlate such firing with what subjects say and do. We must see, for instance, that certain synapse firings reliably correlate with action that helps others, or with people telling us of their altruistic motivation. Without evidence of that connection, brain scans could never tell us whether their subjects were motivated by self-interest or altruism.

..........
Ethical Egoism

Psychological egoism tells us that our only motivation is self-interest. If this is true, then the only motivation we can have is to make ourselves better off. This is not an ethical view, because it says nothing about what is right or wrong. It just claims to tell us about how we actually behave.

But as we have seen, if psychological egoism is true, then there is little point in fashioning a morality that demands self-sacrifice. If all we can be is self-interested, then any morality that required us to be altruistic would be dead in the water. Morality does not demand the impossible. If psychological egoism is true, then all morality could possibly do is to ask that we look out for ourselves.

This is precisely the advice of **ethical egoism**. Unlike the psychological version, ethical egoism really is a moral theory. It tells us about what we are morally required and forbidden to do. Specifically, it says that there is one ultimate moral duty—to improve your own well-being as best you can. Whenever you fail to achieve this goal, you are behaving immorally.

Many ethical egoists have been psychological egoists as well. But this isn't required. You might reject psychological egoism, and so think that we *can* be altruistic, while denying that we *should be* directly concerned with the well-being of others. Ethical egoism gains strength from psychological egoism, but this is not its only source of support. As we will see, there are several arguments for thinking that (surprisingly) ethics is all about getting ahead in the world and making ourselves as well off as we can be.

Why Be Moral?

Imagine that you are a stock broker with inside information about an upcoming corporate takeover. Once the takeover is announced, the company's stock will soar. If you were to purchase a large amount of that stock right now, you'd stand to make millions. Should you do it?

Suppose that you are seriously thinking about this. You've carefully calculated the chances of getting caught, and have decided that the risk is minimal. Further, even if you are caught, the penalty is slight enough as to make the risk worth taking. Still, you are sure that such behavior is immoral. Brokers are able to gain privileged information only on the condition that they not take unfair advantage of it. You'd be doing just that. Yet since the potential gain is so great, why let morality stand in the way?

There are countless scenarios that pose the same problem. Imagine cheating on an exam to increase the chances of landing an excellent job. Or lying to an investigator to avoid prosecution. Or doing some creative bookkeeping to minimize a tax burden. Or spreading malicious gossip in order to undermine a competitor.

In a perfect world, virtue would always be rewarded and vice would never flourish. But what should we do, here in our imperfect world, when immorality promises great rewards? What to do when moral behavior is met with ridicule, a prison term, or a bullet? It is easy when morality and self-interest give the same advice. But what if they don't?

If ethical egoism is true, there are no such cases. This sort of egoism claims that *actions are morally right just because they best promote one's self-interest*. In this view, conflicts between self-interest and morality are impossible, because our fundamental moral duty is to maximize self-interest. If, among all of the options available to you, there is one that will serve you best, then that is the option that morality requires.

It may seem that there is nothing "ethical" at all about such a theory. Consider: if people can best improve their lot in life by secretly killing political opponents, stealing from the weak, or humiliating their employees, then ethical egoism imposes a moral duty to do such things. But how could a plausible ethical theory do that? Indeed, it seems that we can use such examples to construct a decisive anti-egoist argument—call it the *Argument from Paradigm Cases*. The argument would go like this:

1. If an ethical theory requires killing, rape, or theft, just because such actions promote self-interest, then that theory cannot be true.

2. Ethical egoism sometimes requires such things, just because they sometimes do promote self-interest.
3. Thus ethical egoism cannot be true.

A paradigm is a model, a very clear example. Some of the examples used in this argument are the clearest cases of immorality we can think of. The underlying thought here is that any ethical theory ought to be able to make sense of these clear cases, and classify them properly. Ethical egoism fails to do that.

There are two replies that ethical egoists (henceforth, just *egoists*) can offer to such an argument. They can reject the first premise. Or they can reject the second.

To reject the second premise, the egoist must argue that murder, rape, and theft can *never* make the perpetrators better off. That's a pretty hard thing to argue. We can't deny that some people definitely want to kill, rape, or steal. Nor can we deny that a criminal lifestyle sometimes brings great pleasure to those involved. But the egoist might insist that getting what you want or being happy is not the ultimate key to well-being. And we have reason to think that this is right.¹ In this view, what is best for you—even if you don't know it, and even if it doesn't get you what you most want—is to live an upright life.

This claim could be true if there is a God who punishes sin with eternal damnation. It could also be true if immoral acts deserve suffering, and a doctrine of karma ensured that, over the course of successive lives, you got what you deserved. But suppose that you don't find such accounts convincing. Are there other ways to defend the idea that killing, theft, betrayal, and assault are never good for the perpetrators?

Plato thought he could do it. He tried to show that those who are unjust are always harmed because of their injustice. He spent a good part of the *Republic* arguing that all-powerful tyrants, though able to control and acquire so much in this world, are still doomed to terrible lives. Even as they tell themselves how much fun they are having, they are actually suffering great fear and insecurity, racked by emotional turmoil. Once we see what an immoral life is really like, apart from the glamorous depictions of the publicists, we will realize that we are far better off being moral.

Plato's argument depends heavily on his claims about the inner lives of immoral people. I think it fair to say that, in this regard, his case is not fully persuasive. Many immoral people are, to be sure, as emotionally unsettled as Plato makes them out to be. But others are able to sleep well

at night, take pride in a job well done (assassination, theft, betrayal), and find friends within a network of like-minded associates. It seems to be a sad fact that crime does sometimes pay. The bad guys sometimes get away with it, enjoy the fruits of their evil labors, and never regret the harm they have caused. The kinds of feelings, experiences, and outlooks that make up a life worth living—those associated with self-esteem and self-respect, satisfaction at valued work, close friendships and loving relations, good health, and an ability satisfy most of one's wants and needs—have all been enjoyed by many immoral people.

Much more, of course, needs to be said about what is truly beneficial for us. Chapters 1–4 cover these issues, but they are hardly the final word on the matter. And since we don't have space here to discuss religious views of the afterlife, we have to admit that the jury is still out on the truth of premise 2. Perhaps it is false—perhaps there is some very powerful argument that can show that bad people *never* profit from their wrongdoing. If that argument can be given, then we can strike down premise 2. And that, of course, would be great news for the egoist, since the fall of premise 2 would spell the defeat of the entire Argument from Paradigm Cases.

But let's assume for the moment that premise 2 is true, and that, as it seems, some cases of immorality really do benefit those who commit them. If that is so, then egoists must take aim at premise 1 of the argument. And that, in fact, is precisely what they will do.

Suppose that we will not always do best for ourselves by following the familiar laundry list of moral rules. Then conventional morality and self-interest will sometimes clash. When that is so, the egoist is forced to reject conventional morality. And since the first premise of the argument is just an expression of conventional morality, the argument begs the question against the egoist.

Think about it this way. Only those who already dislike egoism would accept the first premise. Premise 1 is not a neutral claim. Rather, it assumes that egoism is mistaken. It assumes, specifically, that self-interest cannot morally justify actions that strike us as very bad. True, actions such as secret killings do seem to be paradigm cases of immorality, and so any moral theory that encouraged such actions would be suspect. But an egoist who accepts premise 2 will claim that our paradigms are mistaken. And they might be. We can't absolutely rule out such a possibility.

Still, we can recall Aristotle's advice, recorded at the end of the previous chapter. We have reason to stick with the appearances, and to take things at face value, until we are given excellent reason for doubt. It *seems*

that killing, rape, and theft are very clear cases of immorality, especially when they are done to serve self-interest. That is just what premise 1 says. Given how appealing that premise is, we are right to insist that egoists provide a compelling counterargument, one that can reveal the error of our popular ways of thinking. Let's now consider two prominent attempts to do just that.

Two Popular Arguments for Ethical Egoism

As we have just seen, ethical egoism can clash very radically with our deepest moral beliefs. This happens in three ways.

First, egoism may *require* some actions that seem grossly immoral. If promoting self-interest requires a knifing in the back, the betrayal of a friendship, or the illegal use of insider information, egoism insists that we do these things.

Second, egoism *forbids* us from doing some actions that seem clearly morally good. Egoists think that any action that involves a genuine self-sacrifice is immoral. Egoism insists that it is wrong to go out of your way to be kind, to keep your promises, or care for your children, if you pass up a chance at personal gain by doing so.

Third, egoism *permits* us to escape some very important moral duties. It seems, for instance, that we each have a duty of easy rescue. If saving someone comes at little or no cost to ourselves, then we must do what we can to help. Yet if your interests are best served by ignoring such victims, then egoism frees you of any duty to help. Indeed, as a general matter, egoism requires that we help others only when we help ourselves in the bargain.

The Self-Reliance Argument

These claims require defense. Here is one that is worth considering—not because it succeeds, but because it is often urged with approval. Call it the *Self-Reliance Argument*:

1. If everyone were to mind his own business, and tend only to his own needs, then everyone would be better off.
2. We ought to do what will make everyone better off.
3. Therefore, we each ought to mind our own business and tend only to our own needs.

There are two problems with this argument. Its first premise is false. And its second premise is one that egoists cannot accept.

The first premise is false: Those who are in need of help would not be better off if others were to neglect them. If you are suffering a heart attack and I know CPR and am the only one able to help, then you are definitely *worse* off, not better, if I decide to leave you alone and go on my way. Nor is complete self-reliance even a good general policy. It may be better if everyone were self-reliant than if everyone were an intrusive busybody. But these are surely not our only two options. There is a middle path that allows a lot of room for self-interest but also demands a degree of self-sacrifice, especially when we can offer great help to others at very little cost to ourselves. Everyone would be better off if people helped others to some extent, rather than if people offered help only when doing so served self-interest.

Further, the argument's emphasis (in premise 2) on our doing what will improve everyone's well-being is not something that the egoist can accept. For ethical egoists, the ultimate moral imperative is to maximize personal benefit. There is no moral duty to improve everyone's lot in life. The egoist allows people to help others, or to have a care for the general good, but only when doing so will maximize their own self-interest. And not otherwise.

The Libertarian Argument

There is another popular argument for severely limiting our duties to others. Call this the *Libertarian Argument*. Libertarians claim that our moral duties have only two sources: consent and reparation. In other words, any duty we have to another person stems either from our voluntarily agreeing to accept that duty (i.e., our consent), or from our having violated someone's rights, and so owing a duty to repair the wrong we have done. But if I do not consent to help other people, and have done them no wrong, then I am free of any duty to offer them assistance. In that case, I am morally entitled to look after my own interests exclusively.

This is a fascinating argument. There is little controversy that our duties can originate as the libertarian suggests. The real question is whether there are sources of duties *other than* consent and reparation. In the example of offering easy rescue, for instance, it seems that the victim's needs, together with my ability to help at little cost, are enough to generate a duty that I help. Consent did not enter the picture—I was morally required to help even if I didn't agree to do so. And I had done the victim no wrong, so reparation was not an issue. The libertarian will deny that someone else's

needs, plus one's own ability to help, are together enough to generate a moral duty of offering assistance. After all, if I need ten thousand dollars for knee surgery, and you are a millionaire who could easily afford to part with that sum, you are not automatically duty-bound to hand me the money.

There is a great deal one might say about the Libertarian Argument. Indeed, I think that it poses one of the most fundamental challenges in political philosophy. Yet we can avoid a look into its details, because even if the argument is sound, it cannot support ethical egoism.

The basic explanation for this is that egoists cannot accept the argument's central claim. Egoists deny that there are two ultimate sources of moral duty (consent and reparation). In fact, egoists deny that *either* of these is a source of moral duty. For them, self-interest is the only source of our moral duties. We must fulfill our voluntary agreements, or repair the damage we've done, only when doing so is in our best interest. *When it is not, we have no moral duty.*

The Libertarian Argument tells us, for instance, that if we agree to volunteer at a local hospital, or consent to the details of a home sale, then we should follow through. Suppose we have no second thoughts about the deals we've made. We really want to volunteer, or to purchase the home. Still, if doing so fails to make us better off, then egoism says that we have no duty to keep our word. Indeed, egoism will *forbid* us from keeping our end of the deal. Libertarians would require that we keep our word. Since egoism and Libertarianism often give such conflicting advice, egoism cannot gain support from Libertarianism.

The Best Argument for Ethical Egoism

Despite their relative popularity, the Self-Reliance and the Libertarian arguments will not succeed in supporting ethical egoism. To locate a stronger foundation for the view, I think we need to focus on one of the perennial questions in ethics, and see how well the egoist can answer it. It is here that egoism finds its strongest support.

The perennial question is the one we began with: Why be moral? Ethical egoism has a watertight answer to this question. We should always be moral because morality always serves self-interest. We all agree that there is good reason to look after ourselves. Since that is so, and since morality (as the egoist sees it) always advises us to protect our interests, there is always good reason to do as morality requires.

We can mold this line of thought into a very powerful argument for ethical egoism. Call this the *Best Argument for Ethical Egoism*. It starts with the popular thought that every moral duty provides an excellent reason to obey it. And it then says that we have reason to do things only if there is something in it for us. It would be irrational, for instance, to sacrifice your own well-being if you got nothing in return for such sacrifice. When you put these two thoughts together, you arrive at the conclusion that whenever we are morally required to do something, doing it must promote self-interest. This is just what ethical egoists believe.

Here is the argument in a nutshell:

1. If you are morally required to do something, then you have good reason to do it.
2. If there is good reason for you to do something, then doing it must advance your interests.
3. Therefore, if you are morally required to do something, then doing it must advance your interests.

The argument is logically valid. Its two premises are each widely accepted. And on the face of it, they are each highly plausible. Consider the first premise. If I am duty-bound to do something, don't I have some good reason to follow through? Perhaps this good reason isn't always decisive—there might sometimes be even better reasons that count against doing my duty. But at the very least, being morally required to (say) keep a promise, or tell the truth, is at least some reason to do such things.

Now consider the second premise. It is hard to doubt that there are compelling reasons to protect your own interests. If some action promises me no gain, but only loss, then what reason is there for me to do it? I'd be irrational to sacrifice my interests without the promise of some compensating benefit. It might be heroic to give up one's life for a stranger or to forfeit one's last real chance at happiness so that others may enjoy life. But reason can't require such sacrifice.

My own view is that the best argument for ethical egoism is ultimately unpersuasive. I think that the first premise is true, but have come to doubt the second premise. Once we examine this theory of reasons with a little more care, it may be less plausible than a first look would suggest.

We can begin to see this by considering two superficially similar claims:

- (A) If an action advances your self-interest, then there is good reason for you to do it.

- (B) If there is good reason for you to do an action, then doing it must advance your self-interest.

Claim A looks pretty good. It tells us that self-interest is always a good reason for doing something. It doesn't state that this is the only reason for acting. And it doesn't state that it is always the best reason. Though A is an extremely attractive claim, it is difficult to explain its plausibility. It seems to me just one of those rock-bottom assumptions we make about the kinds of reasons we have. If someone were to deny it, for instance, it isn't clear what we could say on its behalf. That doesn't show it to be implausible. Not everything can be explained. It may be that A is more basic than any other claim we might use to defend it.

This isn't the case for B, which should look familiar. It is the same claim as the Best Argument's premise 2. And it seems that there are many counterexamples to it. Cases of easy rescue provide the most compelling ones. If I see a terrible traffic accident occur, and have a cell phone with me, I have a reason to dial 911. I have that reason even if making the call will gain me nothing. If I see a window washer high up on a ladder and see that my walking companion is about to accidentally bump into the ladder and send the man to his death, I have a reason to tell my friend, or take her by the arm and redirect her steps. I have that reason even if there is no benefit to me of doing so.

Defenders of premise 2 cannot accept this, and that seems a strike against their view. We can allow that promoting self-interest is a good reason for doing things. But why, as they claim, is it the *only* good reason? To insist that your own interests are all-important, and that the interests of others provide no reasons at all (except insofar as helping them helps you), is a claim that requires substantial defense. I am not sure what that defense would be.

In the absence of such a defense, we have some basis for thinking that premise 2 is false. And if it is false, then the Best Argument for Ethical Egoism is unsound, since that argument relies on it. This does not by itself show that ethical egoism is false. But if we can combine the absence of a strong argument for ethical egoism with the presence of strong reasons to oppose it, then this does leave us with a compelling case against ethical egoism.

Three Problems for Ethical Egoism

The three most serious problems for ethical egoism are (1) that it violates some of the deepest and most central moral beliefs we have; (2) that it

cannot allow for the existence of moral rights; and (3) that it arbitrarily assigns self-interest complete priority over the interests of others.

Egoism Violates Core Moral Beliefs

We have already mentioned the first criticism, and have allowed that it can't refute egoism all by itself. Still, if a theory deeply violates common sense, then we are justified in rejecting it unless a compelling argument can be made on its behalf. Ethical egoism does run strongly counter to common sense, since it imposes a moral duty to kill, rape, torture or humiliate whenever doing so best serves self-interest. It allows us to ignore the vital interests of others, even if it costs us next to nothing to attend to them. We have failed to locate a compelling argument for ethical egoism. Until we do, we are therefore justified in trusting our moral paradigms, which cast doubt on egoism.

Egoism Cannot Allow for the Existence of Moral Rights

It isn't clear that egoism can make sense of moral rights. These are moral claims that give a person control over certain aspects of her life, even if it is to another person's advantage to ignore such moral claims. If I have a right to be free of physical attack, for instance, then it is wrong of anyone else to beat me to a pulp, even if doing so somehow makes them better off. Egoists can offer no one a guarantee against this (or any other) sort of personal violation.

Suppose that it is in your best interest to kill me, or to deny me freedom of speech, or to take what I own. If egoism is true, then you are morally permitted—indeed, morally required—to do these things. And if you are morally allowed or required to do them, then it is hard to see how I could have a right that you not do them. My right to life or to free speech is worthless if it may be infringed whenever it is in someone else's interests to do so.

Egoists can reply that they grant each person the right to pursue self-interest. And that is true. But this is not the sort of right that offers any protection. It simply says that we are free to pursue our own self-interest. It does *not* say that everyone else must keep their hands off when I am doing this. Whenever my self-interest conflicts with yours, I have a duty to harm you, since that will benefit me. And you have a duty to harm me for the same reason.

Thus the right to pursue self-interest offers us no real moral protection. And egoists cannot make sense of any other rights we may have, since

such rights might only stand in the way of people fulfilling their basic moral requirement—that of promoting self-interest, even if it comes at another person's expense.

Egoism Arbitrarily Assigns Priority to Self-Interest

Now we arrive at the deepest threat to ethical egoism. This theory tells us that we should give exclusive priority to the interests of a single person (oneself) over the interests of everyone else. If that were really so, then there would have to be some relevant difference that could justify this different treatment. But what could that be? My basic needs—for food, shelter, physical security, good health, etc.—are shared by nearly every human being. I am unique, of course, but then so is everyone else. I have my special talents, but again, so does everyone else.

If ethical egoism is to be vindicated, its defenders need to explain why we are allowed to entirely discount the interests of others, even though their fundamental interests may be identical to our own. I think that egoists can, in fact, make a partial reply here.

Suppose that my leg has been wounded in a hunting accident. I can still drive to the hospital and get it cared for. I know that doing so is going to cost me almost all of the money I have. Once I arrive at the hospital, I see another victim of a hunting accident, with an injury very like my own. He is obviously poor, needs the surgery as badly as I do, but lacks the money to pay for it. Most of us agree that it would be acceptable if I used my money to pay for my own surgery, rather than forgo it and pay for his. But why?

To make the case harder, assume that this other accident victim is as nice a guy as I am, is just as smart and community-minded as I am, etc. Still, I am allowed to give myself preferential treatment here. Our cases appear alike, and yet common sense allows me to give myself priority over another person who is my equal in every relevant respect.

Perhaps common sense is just mistaken in allowing people to give some priority to their own needs over those of others. If it is, then ethical egoism is certainly false. But suppose that common sense has it right, so that in the hunting example, for instance, I am allowed to spend the money on myself, rather than on a stranger. If so, then morality does grant individuals some extra consideration when determining their own fate. When things are equal, we are allowed to tip the scales in our own favor.

I do not know how to explain this. Like the claim that we have some reason to promote self-interest, this principle seems to be a basic,

fundamental element in any plausible ethical view. Yet even if we accept that people are morally allowed to give themselves some priority, that does not mean that they are allowed to give themselves *complete* priority over others. Ethical egoism claims that the interests of others, considered in themselves, *count for nothing*. We are to treat our own interests as the only thing of importance.

Ethical egoism tell us that even when the stakes are very great for others, and very small for ourselves, we are entitled to ignore the needs, wants and interests of those who are our equals in all relevant respects. It completely denies the moral importance of other people or other things (such as the environment), except insofar as they can help us to benefit ourselves. That is a kind of deep bias that requires substantial defense. It isn't clear what that defense would be.

Conclusion

A theory that requires us only to maximize self-interest (even if doing so means killing or torturing others), and one that frees us of any direct duties to help others in need, is one that requires some very strong argument. We have not met with such an argument. Until we do, we are right to doubt that morality is as limited as ethical egoism makes out.

If doubts about ethical egoism are on target, then morality does indeed assign some intrinsic value to the interests of others. But how far should this be taken? Must we sacrifice *everything* for others—even for strangers? Or is there some middle ground, a principled way to balance our own needs with everyone else's?

The next theory we consider, consequentialism, will begin the process of shifting the moral focus away from ourselves and toward other people. As we'll see, consequentialists don't deny the importance of self-interest. They just deny that your interests are any more important than anyone else's. This move to impartiality is very important in moral thinking. Let's see how far it can take us.

Notes

1. Recall the criticisms of hedonism and the desire satisfaction theory in chapters 2 and 4.

Consequentialism: Its Nature and Attractions

.....
Do all the good you can, by all the means you can, in all the ways you
can, in all the places you can, at all the times you can, to all the people
you can, as long as ever you can.
.....

Thus did John Wesley (1703–1791), English religious thinker and founder of the Methodist Church, sum up his philosophy of life. Our business here on earth is to do good—as much and as often as we can. Our efforts must extend beyond ourselves, to all whom we might help. Benevolence should be our guiding aim, and a life of altruism and good works should be the record we leave behind.

Wesley's pithy formula can be condensed even further, to the motto that defines the **consequentialist** outlook: Do as much good as you can. It is a very attractive picture of the moral life. It requires us to move beyond egoistic concerns, and to focus on improving the lives of others, as well as our own. We must make the world the best place it can be. Can anyone seriously argue with such a claim?

Some very fine philosophers have thought it impossible to reject this view. G. E. Moore (1873–1958), an English thinker famous in his day for his rigorous thought, declared it plain that what is right is whatever produces the most good. If you have two choices of what to do, and one brings less good than the other, that choice can't possibly be the right one to make. *Acts are morally right just because they maximize the amount of*

goodness in the world. Moore thought that those who failed to accept this simply didn't know what they were talking about.¹

It doesn't take much acquaintance with our subject to know that one philosopher's truism is the next philosopher's bunk. Certainly, any plausible moral theory will insist on the central importance of doing good. We commonly justify our actions by pointing to the good they did, and criticize actions by showing that they caused unnecessary harm. That makes perfect sense if consequentialism is correct.

Before we get to the details of the consequentialist picture, let's briefly consider a contemporary conundrum—the morality of capital punishment—that can illuminate the basic nature of this ethical outlook. There are many views about the morality of the death penalty, but most of them can be sorted into two large groups. In the first, consequentialist camp, people insist that such punishment is justified only if it improves our lives. It will have to decrease crime, increase security, and expand respect for human life. If capital punishment is to be justified, we must show that we will be better off with it than without it. We must look to the future, and ask three things: What are the benefits of executing criminals? What are the drawbacks? Which policy would yield the greatest cost-benefit ratio? Survey all of the available options. Whichever policy is **optimific**—i.e., such as to yield the greatest balance of benefits over drawbacks—is the one that morality requires.

A second group asks not about what the future will hold, but rather about what the past requires of us. Specifically, the focus is on whether certain people deserve to be killed for the crimes they have committed. In this line of thinking, even if the death penalty is extremely costly, fails to prevent crime, and perhaps even increases the crime rate by making the population more hardhearted, we should still impose it if we could show that criminals deserve to be executed. Before we consider making people happy, or reducing their misery, we must first satisfy the demands of justice. If giving out just deserts happens to reduce crime, so much the better. But if it doesn't, we should still execute the murderers among us if they deserve it.

My aim here isn't to try to resolve this thorny issue, but rather to reveal what is distinctive about the consequentialist approach to ethics. Consequentialists are those who encourage us not to cry over spilt milk. They direct our attention to the future, not the past. They ask us to look at the *consequences* of our actions or policies—hence the name of their theory. For them, the ends justify the means, so long as the ends are good enough.

When we want to know whether our plans are in line with morality, we will ask about results. Did my action achieve better results than any other I could have done? If so, then my action was morally required. If not, then it wasn't.

The Nature of Consequentialism

Its Structure

Consequentialism says that *an action is morally required just because it produces the best results* (i.e., is *optimific*). But how can we determine whether an act produces the best possible results? It won't always be an easy thing to do in practice. But in theory, it's pretty straightforward. There are five steps to this process.

1. First, identify what is worth having for its own sake. Familiar candidates include happiness, knowledge, virtue and friendship. Forget about any side-effects its pursuit might bring, and focus just on the thing itself. Then ask: Is there anything valuable about this thing? If the answer is yes, then it is intrinsically good—valuable in and of itself.²
2. Next, identify what is intrinsically bad, i.e., bad all by itself, quite apart from any regrettable results it may cause. Examples might include physical pain, mental anguish, sadistic impulses, and the betrayal of innocents.
3. Then determine all of your options. What choices do you have? Which actions are open to you at the moment?
4. Then ask, of each option, what the value of its results would be. Think in particular of how much intrinsic value will be produced. Determine what the intrinsically bad results will be, if any.
5. Finally, pick the one that yields the best balance—the highest ratio of good to bad results. That is the *optimific* choice. That is your moral duty. Doing anything else—failing to strike the greatest balance of good over bad—is immoral.

We can develop dozens of different versions of consequentialism, depending on which things we regard as intrinsically valuable. We could say, for instance, that acts are right just in case they yield the greatest improvement in environmental health, or best advance the cause of world peace, or do more than any other action to increase the amount of knowledge in the world. Each of these is a version of consequentialism.

Thus consequentialism isn't just a single theory, but is rather a family of theories, united by their agreement that results are what matter in ethics. We can't discuss every member of the family here, so I will restrict my attention, for the most part, to its most prominent version—**act utilitarianism**. Act utilitarians endorse the idea that well-being is the only thing that is intrinsically valuable. And faring poorly is the only thing that is intrinsically bad. Their view, then, is that *an action is morally right if and only if it does more to improve overall well-being than any other action you could have done in the circumstances*. Philosophers call this ultimate moral standard the **principle of utility**.

Some utilitarians have been hedonists; others, desire satisfaction theorists; still others have defended the idea that there are a variety of things that directly contribute to our welfare. At present, we don't need to take a stand on this matter.³ The important point is that, according to utilitarianism, acts are right just in case they maximize the overall amount of well-being in the world.

When there are attractions and difficulties that are specific to act utilitarianism, as opposed to other versions of consequentialism, I note them. But for the most part, we can get a good sense of the consequentialist ethic just by focusing on act utilitarianism.

Maximizing Goodness

If act utilitarianism (henceforth, just *utilitarianism*) is correct, then we are duty-bound to maximize well-being. But what, exactly, does this mean? John Stuart Mill (1806–1873), one of the greatest utilitarians, famously summarized the utilitarian outlook by saying that it required us to create *the greatest good for the greatest number*. But this popular catchphrase overlooks an important nuance, one that Mill was well aware of.

Mill was a hedonist, who believed that only happiness was intrinsically valuable and only misery was intrinsically bad. Let's just assume the truth of hedonism for the moment, to better appreciate what maximizing goodness requires. If we combine utilitarianism and hedonism, we get this ultimate moral principle: produce the greatest *overall balance* of happiness over misery. There are two common misunderstandings of this principle. Let's clear them out of the way first, so that we can see what Mill really had in mind.

First misunderstanding: in choosing between acts that benefit people, we must benefit the greatest number of people. Mill rejects this.

Suppose we have to choose between an act that benefits more people, and one that benefits fewer. Mill's principle does *not* say that we automatically have to pick the first option. That's because the benefit to the majority may be very small, whereas the benefit to the minority may be very large. Suppose the legislature has a surplus this year. They could spend that money giving 90 percent of the citizens a fifty-dollar gas coupon. Or they could spend that money trying to prevent homelessness and starvation among the poorest 10 percent. On a variety of plausible scenarios, much more good will be done by giving the money to the poor. That benefits fewer people, but does more good. If that is so, then the money must be directed to combat poverty.

Second misunderstanding: we must always choose that action that creates the greatest amount of happiness. Mill also rejects this.

Suppose we have to choose between two plans. The first creates more happiness than the second. Mill's principle does *not* say that we automatically have to favor the first plan. That's because the first plan may also create a huge amount of misery, while the second plan creates very little. Suppose, for example, that you were a Roman emperor intent on pleasing your public. Your choice: offer gladiator contests, or present a series of grand athletic competitions. Suppose that most people would strongly prefer to watch the gladiators. That creates a greater amount of happiness than the athletic games would. But imagine, reasonably, that this choice also creates much more misery than the athletic games would do. It might be that the greatest *net balance* of happiness over misery is created by the athletic competitions, even though more *total* happiness would have been created by offering a chance to see men killing one another.

The correct interpretation: Utilitarians tell us to do what brings about the best overall situation, by striking the greatest net balance of happiness over unhappiness. So out with the gladiators and in with the athletes.

Moral Knowledge

Utilitarians make the rightness of an action depend on all of its results, no matter how long after the action they occur. There is no statute of limitations on counting consequences. Sometimes the results of one's actions do not last more than a day. In other cases, they may outlast the initial event by decades or centuries. Just think of Booth's assassination of Lincoln or the Romans' crucifixion of Jesus. It seems plausible to say that these actions continue to have effects even today. If that is so, then we must count these effects when determining the morality of these historic events.

This raises an immediate worry. If the rightness of an action depends on all of its results, and these haven't yet occurred, then how can we know whether the action is the right one to do? Utilitarians are split on how best to answer this question. The first group thinks that the morality of actions depends on their *actual* results; the second group thinks that it depends instead on their *expected* results. This makes a big difference in the utilitarian account of how we gain moral knowledge.

The standard view is the one presented earlier. Right actions are those that actually bring about the best possible results. But since these are still in the future when the action is performed, we can never be completely sure that what we are about to do is the right thing. Yet every moral theory has to allow for some degree of moral ignorance—cases where we don't know what is right or wrong. We are not morally infallible. Many utilitarians explain this by saying that since we don't have a crystal ball, and can't perfectly predict the future, we can never be certain in advance that our actions are morally right.

Still, there is a huge amount of past experience that we can call upon. For instance, we can be reasonably sure that if we were to pick up a gun and shoot a perfect stranger at point blank range, then the results would not be optimific. We can't be absolutely certain of this—that stranger could be the next Stalin or bin Laden, for instance. But we should not expect certainty when it comes to moral questions.

Other utilitarians are uncomfortable with this. They want to make moral knowledge easier to get. As they see it, we should be able to know the morality of our acts when we perform them, rather than having to wait until all of the results are in. For, in some cases, that means *never* knowing the morality of our actions, since they may have unexpected results that occur even after our deaths.

For utilitarians in this second camp, the solution to this worry is to make the rightness of actions depend not on their actual results, but on their expected results. In this view, *acts are morally required just because they are reasonably expected to be optimific*. Since we are often in a position to make reasonable predictions about the outcomes of our actions, we can often know, prior to performing them, whether they are morally right or wrong.

Actual versus Expected Results

Suppose that, with the best possible motives, I take an aged stranger's arm and help him across the street. As we cross, a reckless driver crashes into my companion, killing him. Had I not offered him my arm, he would

still be at the curb, a bit delayed, but unharmed. If the morality of an act depends on its actual results, then, from a utilitarian perspective, I did something immoral. A different option (refusing to help the man across the street) would have had better results.

If that sounds harsh, there is an easy remedy. If we evaluate actions based on their *expected* results, then morally speaking, I am in the clear. That's because anyone in my shoes could reasonably expect that helping the man across the street would have excellent results.

Interestingly, most utilitarians reject the amended view that makes rightness a matter of expected, rather than actual, results. The amended version expands possibilities for moral knowledge, and does not condemn actions that are reasonably expected to be optimific. But it has two problems that have lost it most of its supporters.

First, it will sometimes *require* actions that turn out to have disastrous results, when other options would have produced much better outcomes. Some acts that we expect to turn out well end up doing great harm, and when that is so, this proposal would still regard the disastrous actions as morally right. Suppose, for instance, that I am a member of a parole board and have excellent evidence of a criminal's repentance. I authorize his release and later read that he went on a killing spree just days afterward. Most utilitarians find it extremely difficult to morally prefer an act that resulted in many innocent deaths over one that would have resulted in none. They would say that my action was the wrong one, though (as we'll shortly see) I am not to be blamed for it, since I had intended to bring about good results with my action.

Second, some actions are expected to turn out badly, but end up with surprisingly good results. If utilitarians make rightness a matter of expected results, then they must *condemn* such actions. But condemning an act with terrific results, especially when its alternatives are either harmful or much less beneficial, is too much for most utilitarians to stomach.

Assessing Actions and Intentions

If we stick to the classic version of utilitarianism, one that judges actions on the basis of their actual results, what should we say about cases where good intentions yield awful results, or bad intentions yield pleasant surprises? Here utilitarians will insist on separating the issues. They will have one standard for evaluating actions, and another for judging the intentions of the people who perform them.

Actions are right provided they are optimific. But intentions are morally good provided that they are reasonably expected to yield good results. So, in the case of my helping the pedestrian, these utilitarians would judge my action wrong, while praising my intentions. We cannot, in such a view, immediately say that I am to blame just because I did something immoral. Nor can we say that I am to be praised just because I did something right.

Imagine a would-be assassin whose bullet misses its target, ricochets, and miraculously hits both the assassin and his accomplice, disabling them, revealing their location, and leading to their quick capture. In that case, we condemn the assassins, but say of their actions that they are right, since they had the best possible results. In this view, *there is no essential connection between the morality of an action, and the morality of the intentions behind it*. If you are trying to do harm, when you had the chance to do good, then you are to blame—even if your action manages to maximize well-being. If you intend a kindness, then you are to be praised, even if, through no fault of your own, your action backfires and only misery results.

Here is the picture thus far. Utilitarians say that an action is morally required if, and only if, it maximizes the overall amount of well-being. Though theorists differ, most claim that only actual results, rather than those we can reasonably expect, are to be tallied in determining whether this balance has been struck. All results count, not just those that occur in the short term. When we fail to maximize the good, we act wrongly, even if we had the best intentions. Though good intentions may earn us praise, the acid test of an action's morality is its results. When we pass up a chance to do an action that would have had better results, we are doing something wrong. Always.

The Attractions of Utilitarianism

Impartiality

Utilitarianism is a doctrine of impartiality, and this is one of its great strengths. It tells us that the welfare of each person is equally morally important. Whether rich or poor, white or black, male or female, religious or not, your well-being is just as serious a moral concern as anyone else's. In the gladiator example, for instance, we didn't have to ask whether the good results would be primarily distributed among the wealthy nobles or the impoverished slaves. It didn't matter. Everyone's well-being counts, and everyone's well-being counts equally.

We nowadays take such equality for granted. But the idea, when advanced by the earliest utilitarians, was initially quite a radical one. Indeed, utilitarians have a long history of challenging conventional moral wisdom. Jeremy Bentham, whose *Principles of Morals and Legislation* (1781) gave the first really sophisticated defense of utilitarianism, was a fervent abolitionist. Bentham's godson, John Stuart Mill, wrote one of the earliest and greatest works on behalf of female equality, *On the Subjection of Women* (1869). Peter Singer, perhaps the most famous living philosopher, continues the utilitarian tradition of seeking to expand our moral concerns with his widely known work on behalf of ethical vegetarianism and against animal experimentation.

According to utilitarians, we develop a truly moral perspective on things only when we advance beyond narrow self-regard and a concern that is limited to family, friends, co-religionists, or fellow citizens. The moral point of view is nothing less than an impartial concern for everyone whose well-being may be affected by our actions.

The Ability to Justify Conventional Moral Wisdom

As the examples of abolitionism, sexual equality, and animal welfare show, utilitarians are not afraid of controversy. That their recommendations sometimes offend popular opinion is not, they think, a strike against their theory. Rather, it is a strike against conventional wisdom.

That said, utilitarians think that most of our deeply held moral beliefs are correct. There is a special reason for this, one that utilitarians regard as a great advantage of their view. As they see it, utilitarianism does a better job than any competing moral theory in justifying our basic moral beliefs.

Consider the things we regard, deep down, as seriously immoral: slavery, rape, humiliating defenseless people, killing innocent victims. Each of these clearly tends to do more harm than good. Utilitarianism condemns such acts. So do we.

Now consider the things we strongly believe to be morally right: helping the poor, keeping promises, telling the truth, bravely facing danger. Such actions are highly beneficial. Utilitarianism commends them. So do we.

Utilitarianism's ability to justify moral common sense extends not only to the actions we regard as clear cases of morality and immorality, but also to paradigm cases of virtues and vices. Compassion, kindness, and benevolence appear on almost anyone's list of central moral virtues.

Utilitarianism can explain why that is so. Such character traits tend to promote our flourishing and to reduce misery.

Greed, malice, vindictiveness, ingratitude—these are bound to make most people's top ten list of moral vices. Utilitarianism easily explains this. When people are motivated by such things, they tend to harm others, and to miss chances to improve people's happiness. Utilitarians thus condemn such traits as vices. So do we.

Our core moral beliefs could be mistaken. But until we see an excellent argument for such skepticism, we are right to decide among competing moral theories by selecting one that justifies our deepest beliefs, and clashes with only a few of them. No moral theory manages to do this perfectly—as we'll see, each of the major views captures some fundamental moral beliefs, at the expense of some others. Still, utilitarianism earns pretty high marks in this department.

Conflict Resolution

One of the most important things an ethical theory can do is to provide advice about how to resolve moral conflicts. Utilitarianism delivers. Its endorsement of just a single ultimate rule—maximize well-being—enables it to offer concrete guidance where it is most needed.

Consider this familiar moral puzzle. I overhear some nasty gossip about my friend. She later asks me whether people have been talking about her behind her back. I know that she is extremely sensitive, and that if I answer honestly it will send her into a downward spiral for several days. I also know that the source of this gossip is someone who actually likes my friend, and was acting impulsively and out of character. She's probably feeling bad about it already, and probably won't repeat this unkindness.

Of course we need to know a lot more facts about the situation before we can be confident about a recommendation, but if we just stick with the details given here, the utilitarian will advise me not to reveal what I have heard. Honesty may be the best *policy*, but that doesn't mean that full disclosure is always called for. When we consider our choices, utilitarians tell us to opt for the one that will increase overall well-being. Telling the truth won't always do that. Utilitarianism thus gives us a clear method for solving cases of moral conflict, and provides us with a criterion that we can use to solve what might otherwise be very puzzling cases.

John Stuart Mill saw this clearly. In a fascinating discussion at the end of his *Utilitarianism* (1861), Mill touts the superiority of the utilitarian

view of justice over its competitors. He does this, in part, through a discussion of taxation. We can all agree that our system of taxation must be just, but that doesn't get us very far. We can make a case that charging everyone an identical amount of money is just, since we are thereby treating everyone alike. Or we can insist that everyone pay the same percentage of his income—again, a form of treating everyone alike. Or we can require a graduated tax, taking a higher percentage from those who can most easily afford it. Which should we do?

Fans of each of these ideas (and others) will offer many arguments on behalf of their position. Sorting it all out could be excruciating, and that is because we have no clear way of applying the test that tells us to adopt just policies. Utilitarianism cuts through the conflicting claims and offers direct advice: maximize well-being. It's not necessarily easy to determine which tax policy will do this, but at least we have a clear direction in which to pursue our question, once we have settled on our theory of the good life. Compare that to an order to do justice. That's not very helpful advice. If utilitarianism is true, we know precisely what factors are going to make actions right and wrong (namely, the amount of happiness and misery they produce, or the amount of desires satisfied, etc.). We can't always say the same of its competitors.

Moral Flexibility

During the winter of 1846–47, members of the Donner party, traveling by horse and coach in search of a pass through California's Sierra Nevada range, found themselves buried in heavy mountain snows. It wasn't long before their supplies of food and fuel ran out. Almost half of the eighty-seven members of their party died that winter. Some of those who survived faced a terrible choice. They could die by starvation, or remain alive by eating the remains of their fellow travelers.

In the choice faced by members of the Donner party, one might regard the prohibition on eating human flesh as **absolute**—not to be violated under any conditions. Even though honoring the rule means losing one's life, some rules are never to be broken.

Utilitarians would disagree. And their disagreement is based not on specific views about cannibalism, but on very general grounds. For utilitarians, no moral rule (other than the principle of utility) is sacrosanct. It is morally okay to violate any rule—even one that prohibits cannibalism, or torture, or the killing of innocents—if doing so will raise overall well-being.

Utilitarians think that it is an excellent thing that so many of us are unwilling even to contemplate such actions. Our reluctance is almost always beneficial. But in unusual situations, we are right to violate taboos or break familiar, deeply held moral rules if we maximize well-being in doing so.

Utilitarianism is thus a doctrine of moral flexibility. Most of us think that moral rules must allow some exceptions. But where to draw the line? How do we know whether to follow a moral rule or to break it? Utilitarianism gives us an answer. Morality is not a free-for-all. It is not a case of "anything goes." We ordinarily do best when we obey the familiar moral rules (don't steal, lie, kill, etc.). But there are times when we must stray from the conventional path in order to improve overall welfare. When we do this, we do right—even if it means breaking the traditional moral rules.

The Scope of the Moral Community

Suppose that a hiker in the woods comes across an injured deer. It is obviously hurt and unable to run away. Imagine that the hiker, rather than returning to camp and reporting this, decides to very slowly carve the deer to pieces, inflicting the greatest possible pain. The hiker doesn't do this to make it easier to eat the deer—he has no intention of doing that anyway. Rather, he does it because he relishes his sense of power, and enjoys seeing the animal suffer under his knife.

Suppose that this man never does anything like that again. It's not that he's sorry for what he did, but he simply never has the chance to repeat himself. Further, he never even considers treating his fellow human beings like that—they're human, after all, and the deer was just a lowly animal.

What this man did was very wrong, and utilitarians have a ready explanation of this. Most competing moral theories do not. Utilitarians argue that animals are members of the **moral community**. This means that their well-being is morally important in and of itself. That is because animals can be well-off. Their lives can go better or worse for them. That is the utilitarian test for membership in the moral community.

To be a member of the moral community is to have intrinsic moral importance. It is to be owed a certain amount of respect. We have duties directly to members of the moral community. Tables and chairs are not members. Neither are cars. We owe them nothing.

The utilitarian extends the benefits of membership in the moral community to all beings that have a level of welfare. This doesn't mean that

we are never permitted to do them harm. There are many cases in which maximizing overall well-being comes at a price. It may be acceptable to conduct certain intensely painful animal experiments, for instance, provided that they bring about very beneficial results. The point here is that, from the utilitarian perspective, we are not allowed to ignore the happiness and the suffering of any being who can experience such things.

Consider the many beings who are thus enfranchised: infants, toddlers, the severely mentally retarded, and almost all nonhuman animals. Most of us believe that we are morally required to treat these beings with respect, and not (just) because doing so will make us more likely to treat one another well. Regardless of the benefits to folks like you and me, these people, and most animals, are owed some degree of kindness just because of their ability to experience pleasure and pain. That is precisely what utilitarians believe.

Notes

1. G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903).
2. For more on intrinsic value, see the discussion in chapter 1, pp. 18–19.
3. See chapters 1–4 for an extended discussion of the nature of well-being.

..........

Consequentialism: Its Difficulties

Measuring Well-Being

According to utilitarianism, we must do what is optimific. We must maximize overall well-being. Thus to know whether an action is morally required, we need to do four things: (1) add up all of the benefits it produces, (2) add up all of the harm it causes, (3) determine the balance, and then (4) see whether the balance is greater than that of any other available action.

There are many complex cases where it is impossible in practice to follow these steps. There is simply too much information to be gotten, and no one is smart enough or has time enough to gather it all. This explains the extent of our moral ignorance, according to utilitarians. But if it is impossible *in principle* to follow these steps, then utilitarianism is sunk.

Some people do think it impossible. They think that there is no way to undertake step 1 or step 2. Since the problems are the same for either step, let's just focus on step 1. There are in fact two worries here. The first one was discussed in the previous chapter (pp. 116–117), so we can be brief. To add up the benefits or harms of an action, we need to take note of all of its results. Yet for some actions, those results will continue to occur for decades, even centuries, after the action took place. That means that we'll never be able to do the needed measurement. And that means that we cannot know whether the actions are right or wrong.

The utilitarian will reply, as before, that every theory must have a story about why we are sometimes morally in the dark. Utilitarianism accounts for moral ignorance by saying that moral knowledge requires knowing all

of an action's results and their value. It's true that we don't always have a grasp of this information. But that just means (according to utilitarians) that we have less moral knowledge than we may have thought.

Let's turn our attention to the second concern. In order to add up all of the good produced by an action, I would need some way to measure it. Well-being would have to come in quantities so that I could measure the amount of one person's welfare, measure another's, and another's, and add them all together. But that seems completely implausible.

This problem could be solved if well-being were simply a matter of the degree to which our desires were satisfied. We could then determine how well off someone was by noting what percentage of his desires were fulfilled. But as we saw in chapter 4, this view has some serious flaws.

Things will be incredibly complex if we opt for a pluralistic view of well-being. Such a view claims that there are many sources of personal welfare. The list differs from philosopher to philosopher, but familiar candidates include knowledge, virtue, love, happiness, and friendship. The idea is that the more you have of these things, the better your life is going for you.

This sort of view creates really hard problems of measurement. How are we to measure degrees of friendship? Of love? Of virtue? And then combine them into some overall measure of personal welfare?

Pluralism raises hard problems even if we think that there are only two components of well-being, rather than many. Suppose, for instance, that happiness and autonomy are the only things that contribute to a good life. If utilitarianism is right, then we must maximize both of these values.

The problem is that it is not always possible to do this. Sometimes we have to choose between increasing happiness and respecting autonomy. Suppose that a patient's family tells his doctor that news of his imminent death will destroy what little happiness he has left to him. They might be right. Still, withholding that information undercuts the patient's autonomy. In such a case, the doctor may either respect autonomy or increase happiness. She can't do both.

There are many cases like this. Should parents allow their teenage daughter to date an older man they know to be abusive? Should families intervene to prevent a relative from joining a cult? Should a superpower undermine the results of a foreign democratic election, if the newly elected government is set on oppressing its citizens and waging war against innocent neighbors? Answering such questions of course requires a great deal more information. But once we have it, we may still face a choice between

respecting autonomy and preventing misery. When that is so, which should we do?

Utilitarianism has no way of answering this question, once we allow that both autonomy and happiness are intrinsically valuable. If happiness were all-important, then we could always give the nod to it. If autonomy were the supreme value, we could insist on respecting it, even at the cost of unhappiness. But if the two are both of ultimate importance, then it is very hard to know what to do if they conflict.

John Stuart Mill himself encountered something like this problem, even though he thought that happiness is the only thing that is intrinsically valuable. He wanted to insulate utilitarianism from the objection that it was a "doctrine of the swine," attractive only to those layabouts who preferred a life of easy pleasure to the more challenging intellectual pleasures. To Bentham's motto "pushpin is as good as poetry" (pushpin is a simple tavern game), Mill retorted that "it is better to be Socrates dissatisfied than a fool satisfied." Bentham urged us to maximize pleasure, no matter its quality. Mill couldn't tolerate that. He insisted that we maximize the *quality* of our pleasures, as well as their quantity.

Leave aside Mill's elitism here, his view that some pleasures are "higher" than others. His deepest problem lies in his demand that we maximize two things (quantity and quality of pleasures), not one. But what if we can't do both? A big bag of chips, a six-pack, and the TV clicker may bring loads of low-level fun. Or I could spend the same five hours trying to reconstruct a difficult philosophical argument. Little pleasure, but very high quality. There is no principled advice that the utilitarian can offer here, so long as both the quantity and the quality of pleasure are of ultimate importance.

Problems of value measurement arise even if there is only a single intrinsic value, such as happiness. After all, happiness comes in many flavors. There is the happiness of momentary elation, of steady contentment, of physical excitement, of mental challenge, of exhausted gratitude. And that is only a small portion of the catalogue. Given these different kinds of happiness, it would be very surprising were there some common measure of happiness that applied to each person.

If my happiness isn't the very same kind as yours, then how can we add up all of the happiness produced by an action and come up with some overall sum? That is the problem of step 1.

I think that utilitarians can offer a partial reply to this problem. They should admit that there is no precise unit that can measure happiness or, more generally, well-being. Nothing can parcel it out the way that hours

can allocate time, dollars money, and miles distance. Still, there are clear cases where some actions create more overall benefit than others. Consider the good produced by a kindly grandmother who takes in her wandering, orphaned grandchildren, and compare that to the benefits of a couple enjoying a friendly game of cards. We can say with confidence that there is more happiness, and a greater improvement in well-being, in the first case than in the second. So long as there are clear cases such as this one, there must be some way to measure well-being, even if it is imprecise.

And we can, of course, say the same thing about harms. There is certainly a harm when a husband curses his wife in front of their children. But there is far more harm when a cholera outbreak hits a refugee camp. We can't say precisely how much more harm there is. But it is clear that there is a greater amount in one case than another. That may be enough to fulfill step 2 of the utilitarian calculus.

Still, there is trouble when it comes to step 3. It asks us to compare the amount of benefits and harms and then, in step 4, to strike the greatest balance. But what does this really mean? I think the best sense to be made of this idea is the one I relied on earlier: The ratio of benefit to harm should be as great as possible. But since neither benefit nor harm really comes in discrete units, we shouldn't take the ratio idea too seriously.

What we can say, on behalf of utilitarianism, is that there will be easy cases, cases where the morality of the action is clear. If, among all of one's choices, there is an act that produces both the most benefit and the least harm, then that is your moral duty. If, among all of your choices, there is no chance of improving well-being, then you must choose the act that will cause the least amount of harm. Of course, we aren't able to precisely quantify harm or benefit, and so which choice produces the most benefit, or the least harm, may not be very clear. But when it is, these guidelines give straightforward moral advice that the utilitarian will endorse.

The problem is that many situations we face in the real world are not like either sort of case I've just described. Whether one act is *optimific* can be quite unclear—not because we are ignorant of all of its results (though we often are), but rather because there is no way to be precise about the amount of benefit and harm it causes. When two or more acts each promise to produce some benefit, at the cost of some harm, there may be no real answer to the question of which action is *optimific*. And this is where we stand in *most* cases. If that is so, then utilitarianism will have lost one of its primary advantages—namely, its ability to give concrete advice about what to do in morally complex situations.

Utilitarianism Is Very Demanding

There are three areas in which utilitarianism seems to demand too much of us mere mortals: deliberation, motivation, and action. Let's consider these in turn.

Deliberation

It might seem that in order to think about how to act, we must first know a huge amount of information. We must know all of the options we face, and their likely results. Then we must determine the overall value of each of our options. And then compare these values to see which yields the optimistic outcome. Sometimes in a matter of seconds.

But that's just impossible. We're not computers. We can't get that kind of information prior to each action (or perhaps any action). Utilitarianism simply demands too much information, and computational skills that no one could possibly possess.

Mill had a ready reply to this objection. Christians can usually know perfectly well what their religion requires of them, without having to reread the entire Bible prior to each action. They rarely need to do fancy calculations to determine their religious duty, because there are centuries of accumulated wisdom that they can rely on in a pinch. So, too, with utilitarianism. We don't need to spend a lot of time thinking about it to know that raping someone will cause more harm than good. In most cases, we can rely on thousands of previous cases to know what is going to be beneficial or harmful.

Of course there are exceptional circumstances, ones where we do have to stop and carefully puzzle out the pros and cons of a given action. But all moral theories allow for these. Furthermore, it simply defeats the purpose of doing good if you spend so much time deliberating. The person likeliest to fulfill the utilitarian goal of improving the world is the one who knows when to stop and think about her choices, and when not to. Too much time pondering one's options can freeze a person, and result in too many wasted opportunities. Far better to act spontaneously in most cases, and reserve time for careful calculation in those rare cases when it is truly needed.

Motivation

Still, doesn't utilitarianism require us to be saints, always on the lookout for chances to do good? Must we *always* be motivated by the desire to do

what is optimific? A plausible moral theory is one that most of us can live by. And asking us always to be prompted by benevolence, never taking more than a moment or two to care about our own needs, means making morality the entire focus of our lives. That simply is too much. A meaningful life is one spent at least partly devoted to yourself and the things that you are deeply interested in. Being always ready to put ourselves and our interests aside, so that we might minimize the world's misery, means passing up a chance at a meaningful life. A moral theory can be disqualified if it sets the bar so high that no one but a saint can meet its standards.

Utilitarians would agree with this. They do *not* believe that we must always be strategizing about how to improve the world. The reason is simple. People motivated in this way usually fail to achieve their goal.

The idea is that those who are always seeking the best outcome are often bound to miss it. This isn't as strange as it sounds. Consider those who seek only their own happiness. Such people rarely achieve their aim. Constantly striving for this goal only makes it more elusive.¹

So, too, with pursuing global happiness. Those who aim only for this tend to be busybodies, cranks, obsessives. (Consider the wonderfully drawn character in novelist Nick Hornby's *How to Be Good*, a decent man whose self-conscious efforts to maximize happiness only bring misery to family and friends.) Almost all of us already have personal relationships, hobbies, and work lives that will prevent our being exclusively focused on improving the world. Trying to reduce your interests and motives to a single one—producing the greatest good for the greatest number—would almost certainly backfire. If that is so, then utilitarians are off the hook. Their theory would not require nonstop benevolence from us.

There is another way to think about these criticisms regarding deliberation and motivation. To understand this new take on the subject, we need to distinguish between a **decision procedure** and a **standard of rightness**. A decision procedure is just what it sounds like—a method for reliably guiding our decisions, so that when we use it well, we make decisions as we ought to. A standard of rightness is a statement of the conditions under which actions are morally right.

Consequentialism is, above all, a standard of rightness. It says that an action is right just because it maximizes the (net) amount of good in the world. This standard is meant to explain precisely why actions are right, when they are.

The key point is that a standard of rightness need not be a good decision procedure. What makes actions right is one thing; the thoughts and

motives that prompt us to act well may be very different. Think of a gymnast during a balance beam routine. The standards of rightness, or excellence, are pretty complicated. There are a lot of subtle ways to go wrong, and many required exercises. The standard of rightness here would include a long checklist of dos and don'ts.

But it would be foolish for a gymnast to be thinking about all of these things during her performance. To achieve her goal of executing her best possible routine, she should have something fairly simple in mind. The decision procedure may be something as uncomplicated as this: breathe deeply and regularly; ignore crowd noise; focus on the middle-distance; be calm. Here, as in many other cases, the standard of rightness is quite different from the decision procedure that will best guide a person to doing well.

According to consequentialists, that's just the way things are when it comes to morality. Most consequentialists think that their standard of rightness—the principle of utility—fails as a decision procedure. Unless we find ourselves in very unusual circumstances, we should *not* be asking ourselves whether the act we are about to do is optimific. The reasons given earlier explain this. Using the principle of utility as a decision procedure would probably *decrease* the amount of good we do in the world. That's because we would probably spend too much time deliberating or second-guessing our motivations, and thereby reducing our chances of doing good. Whenever that is so, utilitarians require that we use something other than the principle of utility to guide our deliberations and motivations.

Action

The last area in which utilitarianism is said to be too demanding is that of action. Even if we needn't always deliberate with an eye to doing what is optimific, and even if we needn't always have a saint's motivations, we really must act so as achieve optimific results. Whenever we fail, we are violating our moral duty, and so behaving immorally. That is bound to strike most people as excessive.

Consider again John Wesley's motto: Do all the good you can, by all the means you can, etc. Such a life would be one of great and constant self-sacrifice. Any time you can do more good for others than you can for yourself, you are required to do so. If you are like most of those reading this book—in no danger of starvation, able to receive decent medical care, able to afford a night out, a new pair of jeans, a vacation every so

often—then utilitarianism calls on you to do a great deal more for others than you are probably doing.

If I have a choice between spending \$1,000 on a beach vacation and sending that money to UNICEF (the United Nations Children's Fund), it's an easy call. UNICEF literature claims that \$1,000 can provide 100 families with a basic water kit for use during emergencies, immunize 1,000 children against polio, or provide enough woolen blankets to cover 250 children during winter-weather emergencies. I'd be unhappy if I had to forgo my beach vacation. But my unhappiness pales in comparison to the suffering of those whose lives could be improved if I spent my \$1,000 on them, rather than myself.

Another way to see this problem is to note that utilitarianism cannot make room for **supererogation**—action that is “above and beyond the call of duty.” Such behavior is admirable and praiseworthy, but is not required. A classic case of supererogation is that of a bystander dashing into a burning building in order to rescue those trapped inside. Utilitarians must deny that even this is a case of supererogation, because they deny that *any* actions are above and beyond the call of duty. Our moral duty is to do the very best we can do. If, among all of the options available to you at the time, dashing into the building is going to minimize harm, then this is what you must do. Attempting the rescue isn't optional. It is your duty.

Utilitarians know that this isn't going to go down very easily, but they reply that morality can indeed be quite demanding. Most of us believe, for instance, that morality can sometimes require us to give up our great enjoyments, and occasionally our lives. Even in more typical situations, our views about how much sacrifice morality requires are colored by how we have been raised. Most of you have been brought up in relatively wealthy societies, where citizens are used to tax rates that leave them a lot of discretionary income. Suppose, instead, that you were raised in a society that celebrated an ethic of giving, one that concentrated primarily on benefiting the poor. In that case, you'd probably think it criminal to give only a small fraction of your money to relieve poverty.

Utilitarians must agree that their view can demand great sacrifices of those who have a lot to give, and frequent and substantial sacrifices from the rest of us. That the implications of a moral theory are burdensome, however, is not a decisive strike against it. That it threatens the status quo and challenges the comforts of the well-off may be a mark of its truth, rather than its falsity.

Impartiality

The impartiality required by utilitarianism really is a substantial benefit of the theory. The happiness of a celebrity or a billionaire is no more important than that of a sanitation worker or a chambermaid. The misery of an impoverished refugee is as morally important as that of a pampered princess. From the moral point of view, everyone counts equally, and no one's interests are more important than anyone else's.

Yet there is also a worrying side to the demand for impartiality. For morality sometimes seems to recommend *partiality*. It seems right, for instance, that I care about my children more than I do my neighbor's children, that I care more for friends than strangers, more for my fellow citizens than those living halfway around the world. And it also seems right to translate my care into action. If I have only a small bit of money, and it could either pay for a dentist to operate on my son's gum disease or relieve the greater suffering of famine victims, most of us will think it at least permissible to pay the dentist. But to do that is to be partial to the interests of my son. Utilitarianism does not allow that. It must reject the idea that a person, just because he is my son, my dear friend, or my fellow citizen, is properly given priority in my decisions about what to do.

Utilitarians can argue that there are still a great many situations in which we should give priority to our near and dear—not because they deserve it or are more important than strangers, but because that is what is most beneficial. They could argue, for instance, that the results of sending my money overseas would actually be worse than relieving my son's suffering. For we must consider all consequences, not just short-term ones. If I were to sacrifice my son's interests so readily, he would feel hurt, and less secure in my love for him. These feelings are bad in themselves and would probably cause further harm in the long run. By contrast, famine victims who don't even know me won't feel slighted by my passing them over so that I can care for my son's needs. So if we take a sufficiently broad view of things, we can see that being partial to the interests of family and friends is usually optimistic after all.

This sort of reasoning is sometimes correct. When all is said and done, we often get better results when focusing on family, friends, and fellow citizens. But not always. After all, in the tale just told, the long-term result of my not sending famine aid is that some people actually die, whereas my son, though in pain and perhaps resentful of my sending the money abroad, would still be very much alive. From an impartial point of view,

the death of famine victims is surely worse than the gum disease afflicting my son. When minimizing harm means giving one's time or money to strangers, utilitarianism requires that we do so. Even if that means forsaking the important needs of friends and family.

This emphasis on impartiality has another unsettling implication. We are to count everyone's well-being equally. But suppose that nearly everyone in a society has a deep-seated prejudice against a small minority group. And suppose, further, that they use this prejudice to rationalize a policy of enslavement. Depending on the circumstances, it could be that utilitarianism *requires* the preservation of slavery in this society.

When deciding the matter, we must take all of the harms to the slaves into account. But we must also consider the benefits to their oppressors. Everyone's interests count equally. Rich or poor, white or black, male or female. So far, so good. But also: ignorant or wise, just or unjust, kind or malicious—everyone's interests count, equally. If enough people are sufficiently mean and ignorant, then utilitarianism can require that we indulge the sufferings they cause. Though such cases are not likely to occur that frequently, they can. And when they do, utilitarianism sides with the oppressors. That is a serious problem for any moral theory.²

No Intrinsic Wrongness (or Rightness)

When considering the sort of case just described, it is tempting to fall back on the following sort of view. Certain types of action (think of torture, rape, enslavement) are intrinsically wrong—i.e., wrong by their very nature. We don't need to check the results of such behavior in order to know that they are immoral. All we need to know is what such actions are truly like. Once we do, we can see that they shouldn't be done.

Utilitarians cannot accept this. For them, the morality of an action always depends on its results. This feature of the theory is precisely what supports its moral flexibility: In principle, any sort of action can be morally right, so long as its outcome is optimific. Even actions that inflict only terrible suffering can be morally right, provided that they manage to prevent even greater suffering. Utilitarianism's moral flexibility comes from its refusal to absolutely prohibit any kind of action.

Suppose that utilitarianism is correct and there is nothing wrong, in and of itself, with killing innocent people. Usually, of course, such killings will not be optimific. But imagine a person who is really badly off and whose future is grim. He wants to live, and enjoys some small pleasures

every now and then. But most of his life is quite bad, and that's not going to change. The world would contain less misery if he were to die. And so we must kill him—provided, of course, that we avoid causing even greater harm by doing so. If we were caught, then we would be sent to jail, our own families would suffer, other vulnerable people would become much more afraid, etc. But if we can kill this man without being caught, while making it look like he died from natural causes, and escape without detection, then utilitarianism tells us that we *must* do so. This is no murder. It is justifiable homicide. That's because this man's life was lowering the overall level of well-being in the world. With him out of the picture, that level has risen.

Utilitarians reject any absolute ban on killing innocents (or torturing them, or stealing from them, etc.). This has a very important implication: any kind of action, no matter how awful, is permitted, provided it is necessary to prevent an even worse outcome.

Thus utilitarians will advise us to engage in various kinds of harmful preemptive actions. These are actions that deliberately inflict harm, so as to prevent even greater harm. Think of the legal defense mounted by various officers in totalitarian regimes—I sent many to their deaths, but if I hadn't done that, a real zealot would have taken my place, and the body count would have been even worse. Some of these defendants are simply lying, and it's easy to see such talk as a lame effort to lessen their guilt. But sometimes these people are telling the truth. If they hadn't taken the job, some true believer would have, with even worse results. If utilitarianism is correct, then those faced with such a choice must overcome any scruples they have, and agree to serve. And once serving, they are given permission by utilitarianism to commit all sorts of horrific actions—provided that less harmful actions would lead to their replacement, and an even greater death toll in the long run.

Utilitarians deny that any type of action is intrinsically wrong. And as they see it, no kind of action is intrinsically right, either. One might think, for instance, that there is something right, in and of itself, about fulfilling a promise, telling the truth, or helping the needy. But utilitarians deny this. The merit of these actions depends entirely on their results.

Suppose I am faced with a choice. I have promised my daughter a special day out, just the two of us. If I keep my promise, we'll both have a lot of fun. But imagine that I could produce just a tiny bit more good by breaking that promise and devoting the afternoon to helping out a stranger. In that case, utilitarianism tells me that I am forbidden to keep my promise. Morally speaking, I have to break it. That isn't an easy thing to accept.

The Problem of Injustice

Perhaps the greatest problem for utilitarianism can be simply put: we must maximize well-being, but doing so will sometimes come at the cost of injustice. Moral theories should not permit, much less require, that we act unjustly. Therefore there is something deeply wrong about utilitarianism.

To do justice is to respect rights; to commit injustice is to violate rights. It is sometimes optimific to violate rights. When it is, utilitarianism requires us to do so.

Consider two examples from wartime: **vicarious punishment** and **exemplary punishment**. Vicarious punishment targets innocents as a way to deter the guilty. Such a tactic often backfires. But it can sometimes be extremely effective. You might stop terrorists from their dirty work by capturing and torturing their relatives. You might prevent guerilla attacks by destroying the entire neighborhood in which they sometimes take cover. Though the torture and deliberate killing of innocent civilians certainly infringes their rights, the utilitarian will require that it be done if it prevents even greater harm.

Exemplary punishment is punishment that "makes an example" of someone. I recently came across a good instance of this in E. L. Doctorow's book *The March*, a fictionalized account of Sherman's 1864 march to the sea. At one point in his campaign, Southern guerrillas captured, tortured, and killed some of Sherman's soldiers while his army was encamped. How to prevent such attacks in future? Sherman had the Southern prisoners of war in his camp brought before him. He selected one at random. Then he had the man publicly shot, and announced that he'd repeat the exercise in the case of future guerrilla attacks. The attacks stopped immediately.

Sherman's command likely saved more lives than it cost, and so, by utilitarian standards, it was the right thing to do. But it came at the expense of the prisoner's moral rights. Shooting that prisoner stopped the guerrilla attacks. But that doesn't mean that the prisoner deserved to die.

Cases of vicarious and exemplary punishment are cases in which people deserve not to be harmed. There are also many examples in which people do deserve some sort of penalty or punishment, but it is not optimific to give them their just deserts. Think of cases in which a student rightly receives a failing grade, and appeals for a better one. Sometimes it really would be most beneficial to give the student the grade he wants, rather than the grade he has earned. Perhaps a job, or a scholarship, or his parents' good opinion of him is on the line. If the benefits outweigh the costs, utilitarianism requires that the professor change the grade.

There are more serious cases. After the Second World War, U.S. officials determined that it was beneficial to allow many Nazi scientists to escape punishment, so long as they agreed to share their weapons intelligence. Prosecutors sometimes let acknowledged murderers go free, if the killers testify against the crime bosses who once hired them. Political leaders with blood on their hands are often allowed to retire peacefully, so as to avoid the civil strife that would result were they prosecuted for their crimes. In each such case, we can minimize harms by failing to give the guilty their just deserts. When that is so, utilitarianism requires that we do just that.

When rights are violated, victims are usually harmed. But not always. A bank manager may embezzle millions, and doctor the records to ensure that his theft is never discovered. He and his family may live like kings, and no one be the worse for it. A thief's last heist may have him stealing a valuable jewel that its owner has come to hate. Suppose the owner felt that he couldn't sell or give the jewel away, but had secretly wanted to be rid of it. The owner's rights have surely been violated, but neither he nor anyone else is the worse for it. In such cases there is only benefit, and little or no harm. Utilitarianism thus approves of such thefts. And yet they were undertaken entirely for selfish reasons, and achieved only selfish gains. These examples, and others offered earlier, should give us pause. A plausible moral theory will have to give justice the importance it deserves. It's not clear that utilitarianism can do this.

Potential Solutions to the Problem of Injustice

For as long as utilitarianism has been around, its fans have had to deal with the objection that it shortchanges justice. They have had ample time to develop replies. There are four that are especially important. We'll consider three of them here, and then devote the next section to the fourth.

Justice Is Also Intrinsically Valuable

The first one is easy. Rather than say that well-being is the only thing of intrinsic value, we should also grant that justice is important in its own right. So we should maximize well-being *and* maximize justice in the world. That will solve the difficulty.

Or will it? It seems, instead, that we just open ourselves up to the problem aired earlier, in which consequentialism loses its ability to give guidance when promoting one value comes at the expense of another.

If we are to maximize happiness and justice, what happens when we can't do both? Which should we give priority to?

We could say: always give priority to justice. But this isn't very plausible. Suppose that there has been gridlock in the state legislature. For months lawmakers have been unable to pass a spending bill. Finally, a compromise package comes to the floor. If it doesn't get passed, there is no telling when another spending package will be voted on. In the meantime, government will shut down, tens of thousands of people will not receive paychecks, medical assistance, or welfare support. Furthermore, the spending bill looks *terrific*. It solves a great number of the state's problems, gives aid to the neediest, and sponsors projects that will do genuine good for most communities. There is only one problem—it includes a clause that unfairly denies a small community the agricultural subsidies that the governor had promised it. Still, given the alternatives, a legislator should definitely vote for the spending bill, even though this means a minor injustice. As a general matter, if the stakes are extremely high, and the injustice very small, then it *may* be right to perpetrate injustice.³

Rather than always giving priority to justice, we might instead always give priority to well-being. But then we are right back to the original theory, and so are stuck with the problem of injustice. What seems right to say is this: Sometimes it's best to prefer well-being to justice. Sometimes not. But without any principle to sort this out, we don't really have a coherent theory at all.

Injustice Is Never Optimific

In the face of this problem, some utilitarians deny that their theory ever requires us to commit injustice. They say that if we carefully consider all of the results of unfair actions, we will see that those actions aren't really optimific after all. A policy of vicarious punishment, for instance, may work in the short run. But it will cause such anger among the target population that an even greater number of them will join up with the terrorists. And that will mean more innocent bloodshed over time.

Such a calculation is certainly true in many cases. But it is unwarranted optimism to suppose that things will always work out so fortunately. Sometimes terror movements really are sapped of support when the surrounding civilian population is forced to take the hit. Sometimes exemplary executions send a message effective enough to prevent an increase in guerrilla violence. Injustice really can sometimes prevent great harm. And it can, on occasion, also yield great benefits. We can't tell the many stories

of the criminals who have gotten away with it, because their happiness depends on their crimes remaining secret. In some of these cases, there is substantial benefit and little or no harm. Utilitarianism must approve of such actions.

Justice Must Sometimes Be Sacrificed

A third utilitarian strategy concedes this point, and allows that well-being and justice sometimes conflict. But when they do, it is justice, and not well-being, that must take a backseat. Justice is only a part, and not the whole, of morality. Of course it is important to respect people's rights, but that is because doing so is usually *optimific*. When it isn't, rights must be sacrificed.

Utilitarians of this stripe know that their recommendations will sometimes clash with conventional wisdom. But as we have seen, this is not a fatal flaw. Received opinion is not the final word in ethics. Utilitarianism began its life as a radical doctrine. That legacy remains.

Utilitarians can claim that our deepest moral convictions, including those that require us to do justice, reflect a utilitarian framework. We are socialized to tell the truth, protect the weak, keep our promises, etc., *because doing so tends to be optimific*. But when it is not, utilitarians ask us to look at morality's ultimate standard, and to set aside our ordinary scruples in favor of the principle of utility.

It is a good thing, from a utilitarian point of view, that we are so reluctant to approve of injustice. But as we saw in the previous chapter, virtuous motivation can sometimes lead to poor results. When it does, utilitarianism condemns the action as immoral. If, for instance, Sherman's squeamishness had prevented his execution of the prisoner, many more innocent lives would have been lost. If innocent lives are so important, then surely (say utilitarians) we ought to save as many as possible. That sometimes requires us to overcome our qualms, and commit acts of injustice.

Most of us agree that justice can sometimes be outweighed by other moral concerns. If, in a previous example, a legislator must authorize a minor injustice in order to pass an immensely beneficial spending bill, then morality gives the go-ahead. If you can administer CPR to a stricken passerby, and so save his life, then it is worth committing a minor injustice to do so. So justice may sometimes be sacrificed. But when? Utilitarians have an answer: whenever the results of doing so are *optimific*. If you don't like that answer, you need to supply a better principle that tells us when injustice is, and is not, permitted.

Rule Consequentialism

There is a moral theory that deserves special mention here, because it promises to handle a number of objections to utilitarianism, while keeping much of its spirit. This is **rule consequentialism**—the view that *an action is morally right just in case it is required by an **optimific social rule***.

An optimific social rule is a rule that meets the following condition: If (nearly) everyone in a society were to accept it, then the results would be optimific, i.e., as good as they could be.

The basic idea is this. Rather than determine an action's morality by asking about its results, we ask instead about whether the action conforms to a moral rule. This is a familiar model in ethics. Most moral theories operate this way. What distinguishes them from one another are their different claims about what makes something a moral rule. Rule consequentialists have a specific view about this. The moral rules are the optimific social rules.

To know whether a rule is an optimific social rule, follow these three steps:

1. Carefully describe the rule.
2. Imagine what a society would be like if just about everyone in it endorsed the rule.
3. Ask this question: Will that society be better off with this rule than with any competing rule?

If the answer to this question is yes, then this rule is an optimific social rule. If the answer is no, then it isn't an optimific social rule, and so is not a genuine moral rule.

Let's make this concrete. Return again to the example in which General Sherman orders a captive publicly shot. Is that a morally acceptable thing to do? According to rule consequentialism, we answer this question by figuring out the moral rules that govern the treatment of prisoners of war. One of these rules might be: do not execute prisoners of war without a fair trial. Another might be: execute prisoners of war if you think that doing so will be highly beneficial. And of course there are many other candidates. But let's just stick with these two.

The rule consequentialist will ask us to imagine societies that are governed by each of these competing rules. Which society would be better off?

That depends, of course, on what makes societies and their citizens better off. Rule consequentialists differ on this question. Some are hedonists; others are desire satisfaction theorists; still others defend different lists of objective goods.

The most prominent contemporary version of rule consequentialism, that offered by the philosopher Brad Hooker,⁴ says that there are two, and only two, things of intrinsic value—happiness and justice. Optimific social rules will be ones that both increase happiness and respect rights.

So, which of the two rules regarding captured prisoners will, over the long run, maximize both happiness and justice? It is very probably the first rule, which requires captors to try their prisoners fairly. A society governed by the second rule would give too much discretion to commanders, who might make serious mistakes about the benefits of killing their captives without due process. There may be isolated cases in which ordering such a killing is optimific. But *as a general policy*, it is probably optimific to forbid such behavior.

If that is so, then rule consequentialists will condemn Sherman's order. And that is surely what justice requires. Rule consequentialism will probably also instruct professors to give their students the grades they deserve, rather than those they would like to have. It will condemn the actions of thieves, even if they don't get caught and their victims suffer in only minor ways.

You see where this is going. When we focus on what is optimific as a general policy, we repeatedly get advice that agrees with our notions of justice. Even rule consequentialists who reject the intrinsic value of justice, and insist that well-being is the only thing of ultimate value, will almost always defend policies that are just. That's because in the long run, and as a general matter, just policies maximize well-being, even if, in isolated cases, just actions do not.

Rule consequentialism also solves other problems that beset act utilitarianism. It supports our belief that morality permits a certain degree of partiality, because policies that allow us to give preference to friends, loved ones, and fellow citizens will very often be highly beneficial.

Rule consequentialism makes it much easier for us to know the right thing to do. Rather than trying to predict the benefits and harms of each available action, and then trying to balance them against one another, we are instructed to follow relatively simple rules. The complex and difficult calculations associated with act utilitarianism are replaced by fairly straightforward rules.

Rule consequentialism can also say that certain actions are simply forbidden, even if they will sometimes achieve very good results. For instance, even if it would be optimific here and now to torture a prisoner, there may well be an optimific rule that forbids political torture. In most cases and over the long run, societies that ban torture may be much better off, in terms of both happiness and justice, than those that allow their officials the chance to torture prisoners and captives. If that is so, then torture is immoral—even if, in unusual cases, it yields real benefits.

So rule consequentialism has a lot going for it. And yet very few philosophers accept it. The reason was given over fifty years ago, by a prominent Australian philosopher, J.J.C. Smart.⁵ In a defense of act utilitarianism, Smart accused rule consequentialists of irrational rule worship. That charge has stuck.

The basic worry is simple. Rule consequentialists demand that we obey moral rules, *even when we know that breaking them would yield better results*. But that is irrational, since in these cases, consequentialists know in advance that their ultimate goal (making the world the best place it can be) will not be fulfilled. It is irrational to knowingly defeat your own goals. Rule consequentialists do this whenever they issue a recommendation that differs from act utilitarianism.

Act utilitarianism demands that we always do what is optimific. So, by definition, whenever rule consequentialists give us different advice, we are required to act in a way that fails to yield the best results. Rule consequentialists would deny military commanders a license to execute the innocent, would forbid torture and embezzlement and vicarious punishment—even when specific instances of such action would be most beneficial. This is self-defeating, since a consequentialist's ultimate aim is to produce the best possible results.

No matter what your ultimate goal is, the rules that *generally* achieve that goal will sometimes fail to do so. If you know that you are in one of those exceptional situations, then why follow the rule? Suppose that justice, not happiness, is the ultimate value. Suppose, too, that justice would be best served if everyone were to follow a certain rule, such as one that prohibits tampering with evidence. But why follow that rule if you know that this time, unusually, breaking the rule will yield the most justice?

If the ultimate purpose of morality is to make the world a better place, then it is irrational to knowingly behave in ways that fail to do this. And yet that is what rule consequentialism sometimes requires. That is why most consequentialists have rejected it.

Conclusion

Consequentialism is a perennial favorite with moral philosophers. Its emphasis on equality and impartiality, its moral flexibility, its inclusion of animals and less-than-fully autonomous human beings within the moral community, its orientation to the future and its emphasis on results have great appeal for many ethical thinkers.

But we have also seen that there are worries for consequentialism, and these are not easily solved. We usually admire impartiality, but sometimes think that partiality is what morality demands. The consequentialist promise of being able to offer concrete advice to solve moral conflicts may be unfulfilled. Consequentialism can require a degree of self-sacrifice that strikes many people as extreme. It sometimes demands that we perform some truly awful actions, so long as they are needed to prevent even greater horrors. And it sometimes calls on us to perpetrate injustice. We considered four solutions to this last problem, none of which seemed wholly satisfying.

It is fitting, then, that we now turn to a competing view that gives pride of place to justice. This is the theory of Immanuel Kant.

Notes

1. Recall the discussion of this point as it relates to the paradox of hedonism in chapter 2, pp. 27–28.
2. Consider this passage from the memoirs of writer Alexander Waugh, who recounts the terrors of corporal punishment at his boarding school: “The monks enjoyed whipping the boys as a release from the constraints of their celibacy, and my father, throughout his life, always claimed that to be beaten was a small sacrifice for a boy and a great treat for a monk.” (This is from Waugh’s book *Fathers and Sons* (Doubleday, 2007).) This remark, no doubt issued tongue-in-cheek, nevertheless illustrates a weakness of utilitarian reasoning. Whether the oppressors are abusive teachers, corrupt politicians, soldiers rampaging through civilian areas, or slave owners, the utilitarian counts their welfare as equal to that of their victims. If no other action would produce as much overall benefit, then the abuse is, by the utilitarian’s lights, morally justified.
3. Everything depends on the details. Much more on this sort of problem can be found in the discussions in chapters 15 and 16.
4. See Brad Hooker, *Ideal Code, Real World* (Oxford University Press, 2000).
5. See J.J.C. Smart, “Extreme and Restricted Utilitarianism,” *Philosophical Quarterly* 6 (1956), pp. 344–354.

The Kantian Perspective: Fairness and Justice

Imagine a person who reasons as follows: I should keep my money rather than pay it out in taxes, because if I keep it, I'll be able to afford a wonderful vacation for myself and my family. And no one is actually going to suffer if I pocket the money, since it's only a few thousand dollars that we're talking about. There's no way that money could bring as much happiness in the government's hands as it could in mine.

Suppose he is right about that. He spends the money on his vacation. He and his family have a terrific time. He is never caught.

Still, he has done something wrong. So has the person who cheats on her exams and gets away with it. So has the person who gleefully speeds down the emergency lane and escapes the traffic jam that the rest of us are stuck in. So has the person whose campaign of dirty tricks has gotten him securely into office.

Despite any good results that may arise from their actions, these people did wrong—or so we think. And the explanation of their immorality is simple. What they did was unfair. They took advantage of the system. They broke the rules that work to everyone's benefit. They violated the rights of others. No matter how much personal gain such actions bring, they are still wrong, because they are unfair and unjust.

Immanuel Kant (1724–1804) thought this way, and was very likely the most brilliant philosopher ever to have done so. He remains perhaps the most important voice of opposition to utilitarianism and its claim that the ultimate point of morality is to improve well-being rather than do justice.

Consistency and Fairness

There is a natural way to understand what is wrong about the actions in the examples just given. In each case, people are making exceptions of themselves. Their success depends on violating rules that most other people are following. This is a kind of inconsistency—of playing by one set of rules while insisting that others obey a different set.

People are inconsistent to the extent that they treat similar cases differently. There's nothing special about the tax cheat or the dirty politician that licenses their actions. They acted as if they had a unique privilege and were exempt from the rules that everyone must follow. But there has to be something unusual about a person, or her situation, in order to gain that sort of privilege. That you can get away with making an exception of yourself doesn't mean that it is right to do so.

Our deep opposition to unfairness, and the corresponding importance we assign to consistency, is shown in some very familiar tests for immorality. The two most popular tests each take the form of a question:

1. What if everyone did that?
2. How would you like it if I did that to you?

When we ask such questions—in the face of a bully, a liar, or a double-crosser—we are trying to get the person to see that he is acting unfairly, making an exception of himself, living by a set of rules that work only because others are not doing what he is doing. These basic moral challenges are designed to reveal the inconsistency, and so the immorality, of that person's behavior.

Consider the first question: What if everyone did that? This question is really shorthand for the following test: *If disastrous results would occur if everyone did X, then X is immoral.* If everyone used the emergency lanes in traffic jams, then ambulances and fire trucks would often fail to provide needed help, leaving many to die. If everyone cheated on their taxes, society would crumble. If every candidate resorted to dirty tricks, then the entire political system would become corrupted. The test works easily and well for these cases.

But the test fails for other cases, and so it cannot serve as a reliable way to learn the morality of actions. Consider a common argument against homosexual sex: If everyone did that, disaster would soon follow, for the human race would quickly die out. Even if this were true, that wouldn't show that homosexual sex is immoral. Why not? Well, consider those who

have decided to remain celibate—perhaps they are priests, or committed lifelong bachelors who believe that one shouldn't have sex without being married. What if everyone did *that*—i.e., refrained from having sex? The same results would follow. But that doesn't show that celibacy is immoral.

The real problem for this test, apart from the fact that it sometimes delivers mistaken verdicts, is that it makes the morality of an action depend on how it is described. Suppose the sexual relations of a gay couple were described as their having consensual, enjoyable sex. In that case, their actions would pass the test. But that undermines the test, because it shows that the test yields contradictory results. The very same action is said to be both morally wrong and morally acceptable, depending only on how it is described. Without any independent guidance on how to select one description over another, this test cannot do the job it was supposed to do—namely, identify which acts are immoral.

What about the other test, the one that asks: How would you like it if I did that to you? This is a direct application of the **golden rule**, which tells you to treat others as you would like to be treated. The golden rule is the classic test of morality. Clearly, it is meant to be a test of consistency. If you wouldn't want to be slandered or exploited, then don't do such things to others. If you do them anyway, you are acting inconsistently, hence unfairly, and therefore immorally.

Getting people to imagine what it would be like to switch places with their intended victims is often a very effective way to convey a moral message. That is why films and literature are often such powerful tools of moral education. But imaginatively filling someone else's shoes, and asking yourself whether you'd accept being treated in a certain way, is actually an unreliable test of morality. The golden rule cannot be correct.

Kant himself identified the basic reason for this. The golden rule makes morality depend on a person's desires. Most of us don't like to be hit. And so the golden rule forbids us from hitting others. So far, so good. But what about masochists who enjoy being hit? The golden rule allows them to go around hitting others. That's bad. The morality of hitting people shouldn't depend on whether you like to take a beating every now and then.

Consider a related problem, that of the fanatic. Fanatics are principled people. It's just that their principles are ones that we find frightening and revolting. Some fanatics are so wedded to their cause, so strong-willed and self-disciplined, that they would accept the suffering that they want to impose on their victims, were the role of victim and persecutor reversed. True, few Nazis, for instance, would really accept a march to the

gas chamber were they to discover their Jewish ancestry. Most Nazis, like most fanatics generally, are opportunists of bad faith, ones with very limited empathy and only a feeble ability to imagine themselves in someone else's place. If roles really were reversed, they'd much more likely beg for mercy and abandon their genocidal principles. But some would not. There are true believers out there who are willing to suffer any harm in the name of their chosen cause. The golden rule licenses their extremism because it makes the morality of an action depend entirely on what you want and what you are willing to put up with.

The golden rule also fails to give us guidance on **self-regarding actions**—i.e., those that concern only oneself. That's not a problem for most people these days, since it's now unusual to think that we owe moral duties to ourselves. But in Kant's time, self-regarding duties were widely endorsed, and many people still think, for instance, that there is something immoral about suicide or about letting one's talents go to waste, even if no one else is harmed in the process.

Because the golden rule sometimes gives the wrong answer to moral questions, it cannot be the ultimate test of morality. Something else must explain why it works, when it does. Kant thought he had the answer.

The Principle of Universalizability

Kant, like most of us, felt the appeal of the two tests just discussed. He agreed that common sense is deeply committed to the importance of fairness and consistency, something that these two tests were trying, but not quite succeeding, in capturing. His aim was to identify the ultimate principle of morality, one that would explain the attraction of the two tests while correcting for their shortcomings.

He thought he had found it in the following standard, the **principle of universalizability**:

An act is morally acceptable if and only if its maxim is universalizable.

To understand what this means, we need to understand two things: what a **maxim** is, and what it is for a maxim to be **universalizable**.

A maxim is simply the principle of action you give yourself when you are about to do something. For instance, if you send a regular check to Oxfam, your maxim might be: contribute fifty dollars per month to Oxfam to help alleviate hunger. A maxim has two parts. It states what you are

about to do, and why you are about to do it. You dictate your own maxims. These are the rules you live by.

Kant thought that every action has a maxim. Of course we don't always formulate these maxims clearly to ourselves prior to acting, but at some level, whenever we act, we intend to do something, and we have a reason for doing it. A maxim is nothing but a record of that intention and its underlying reason. Maxims are what we cite when we try to explain to others why we act as we do.

If we lack a maxim, then we aren't really acting at all. We could be moving our bodies, as we do when we sneeze or roll across the bed in our sleep. But the absence of a maxim in these cases shows that these are mere bodily movements, rather than genuine actions.

Kant thought that an action's rightness depends on its maxim. And this leads directly to a very important implication. For Kant, the morality of our actions has nothing to do with results. It has everything to do with our intentions and reasons for action, those that are embedded within the principles we live by. This is a clear break with consequentialism.

Indeed, we can imagine two people doing the same thing, but for different reasons. That means that they will have different maxims. And even if their actions bring about identical results, one of the actions may be right and the other wrong, since only one of the maxims may be morally acceptable. This is something that act consequentialists cannot accept.

It might be, for instance, that I keep my promise to you because I think it's right to do so. But I might also keep my promise to you because I want you to develop such a trust in me that you leave me your fortune in your will. Assume that the way I keep my promise in both cases is the same. And assume that the results are the same in both cases as well. Then the utilitarian thinks that the morality of my action cannot change between the two cases. But since my maxim is different in these cases, Kant thinks that the morality of these two actions might be different. It all depends, as we'll shortly see, on whether these maxims are universalizable.

Many people agree with Kant's view that the morality of our actions depends not on their results, but on our maxims. For this supports our thought that those who set out to do evil are acting immorally, even if, through sheer chance, they manage to do good. It also justifies the claim that people who live by noble principles are acting morally, even when some unforeseeable accident intervenes, and their action brings only bad results.

Kant had a deep reason for making the morality of an action depend on its maxim, rather than its results. That reason (discussed in detail in the next chapter) is this: it is crucial that the morality of our actions depends entirely on what is within our control. We *can* control which maxims will govern our actions. We decide for ourselves what we intend to do. Even in cases where my options are severely limited, as when a thug has a gun at my head, it is up to me to decide which choice to make.

By contrast, the results of our actions are often out of our hands. We can't always control them. And it is unfair to assign credit or blame for things we can't control. That is why we have an insanity defense. That is why we don't prosecute animals for the damage they sometimes cause. That is why we don't condemn infants for any harm they do.

So the morality of actions depends on their maxims. But how, precisely? Not every maxim is going to be a good one. We need a way to sort out the good maxims from the bad. That's where universalizability comes in.

How can we tell whether a maxim is universalizable? Here is a three-part test:

1. Formulate your maxim clearly—state what you intend to do, and why you intend to do it.
2. Imagine a world in which everyone supports and acts on your maxim.
3. Then ask: Can the goal of my action be achieved in such a world?

If the answer to this last question is *yes*, then the maxim is universalizable, and the action is morally acceptable. If the answer is *no*, then the maxim is not universalizable, and the action it calls for is immoral.

This should strike a familiar note. The test of a maxim's universalizability clearly echoes the rule consequentialist's test for optimific social rules (see the previous chapter), and the *What if everyone did that?* test. Indeed, Kant has us ask a version of that question in the second step of this three-part test. But unlike these other tests, Kant doesn't ask about whether people would be much better off in the imagined world, or about whether disaster would strike there. Instead, he asks about whether we could achieve our own goals in that world. But what is so important about that?

The importance, for Kant, is that this three-part test serves as the real way to determine whether we are being consistent and fair. If our maxim is universalizable, then we are pursuing actions for reasons that everyone

could stand behind. We are not making exceptions of ourselves. Our goals are ones that everyone *could* support, even if, in the real world, some are dead set against them. We are asking whether our aims could be achieved if everyone shared them. If they can be, this shows that we are living by fair rules. Were we making an exception of ourselves, our maxims wouldn't be universalizable.

Consider the tax cheat again. The only reason he can get what he is aiming for (a lovely vacation) is because enough others are not adopting his maxim. The same goes for the careless driver who speeds down the emergency lane.

Kant sought to make this point with an example of his own. (I am embellishing a bit, but the essence of the example is Kant's.) Suppose that I am a compulsive gambler who is constantly in debt. One night I go to the tables to recoup my losses, only to dig myself further in the hole. The casino boss is having his men drop by tonight to collect. I can either pay them or have my kneecaps broken. I know which one I'd prefer.

The problem is, I don't have the money, the bank won't lend me any more, and I don't have anyone to turn to but you. (No one else trusts me any longer to repay my debts). Since you are aware of my reputation, I know that the only way to get the money is by lying to you. So, I beg and plead and promise you, by all I hold dear, that I will repay you—all the while having no intention of doing so. I have just made what Kant calls a *lying promise*.

It seems clear that what I am about to do is immoral. And that is true even if, through a minor miracle, I then feel so much guilt that I repent of my ways, transform myself, and make the lie turn out for the best in the end. The morality of the action doesn't depend on its results, but on its maxim. And my maxim here is not universalizable. So my action is immoral, as Kant says, and as we believe.

Here's why. Suppose my maxim is: lie to a friend, in order to escape from being hurt. And suppose everyone acts on this maxim. They lie whenever they think that it is necessary to avoid some personal harm. In that situation, no one would trust the promises of others. And without that trust, people could not achieve the goals they are aiming for with their promises. In a world where no one believed the promises of others, I'd never be able to get money from you with my promise. And so the purpose of my promise would be defeated. And so my maxim is not universalizable. I am making an exception of myself, and am treating you unfairly. My action is therefore immoral.

Morality and Rationality

Kant claimed that when we act on a maxim that can't be universalized, we are contradicting ourselves. We are being inconsistent. We are assuming that it is acceptable to act in a certain way, even though our purposes could not be achieved if others acted in that very same way. When we make an exception of ourselves, we are acting as if we were more important than anyone else, and going on as if we were exempt from rules that others must obey. But we are not more important than others, and we are not exempt from these requirements.

It follows that when we behave immorally, we are reasoning badly. We are making mistaken assumptions—that we are more important than other people, that the rules applying to them do not apply to us. Those mistakes, and the inconsistent, contradictory reasoning behind them, show that *immoral conduct is irrational*.

That is a very striking claim, and one that most of us hope is true. We want to be able to convict rapists or terrorists of irrationality, of ignoring their strongest reasons. We want to be able to truthfully say that there were excellent reasons for them to do good and to avoid evil. Kant believed that we could do this.

But how could Kant be right? Consider the ruthless contract killer who knows precisely what he wants, knows exactly how to get it, and executes his plan without fail. Morality doesn't enter into his calculations. He knows that what he is doing is immoral, but that doesn't faze him. It seems that such a person is reasoning flawlessly. How could we convict him of irrationality?

Let's call this the *Amoralist's Challenge*. The **amoralist** is someone who believes in right and wrong but doesn't care about morality at all. The amoralist has the same attitude to moral rules as I do to the rules of professional cricket—yes, they really exist, but they have no bearing on my life at all. Obedience to these rules is completely optional. If I am interested in playing the game, then I'll follow the rules. If not, then there is no reason to do so.

The Amoralist's Challenge supports this view in the following way:

1. People have a reason to do something only if doing it will get them what they care about.
2. Doing their moral duty sometimes fails to get people what they care about.
3. Therefore, people sometimes lack any reason to do their moral duty.

4. If people lack any reason to do their moral duty, then violating their moral duty can be perfectly rational.
5. Therefore, it can be perfectly rational for people to violate their moral duty.

The success of this argument would undermine the thought that morality, all by itself, supplies us with good reason to do as it says. It would also refute Kant's claim that immoral actions are always irrational.

Kant thought that you act irrationally when you act contrary to your strongest reasons. And he thought that when moral reasons apply to a given situation, they are always the strongest reasons. Moral reasons defeat the importance of any other kind of consideration. If morality requires you to do something, then that is what you must do—even if you don't want to do it, even if you'll suffer for doing it, and even if the results of doing it are generally disastrous. When you act immorally, you are acting irrationally.

Kant admits that the ruthless contract killer, like so many other successful criminals, did, in a sense, reason perfectly well. He followed what Kant called **hypothetical imperatives**. Specifically, these imperatives (commands) are commands of reason. They command us to do whatever is needed in order to get what we care about. Hypothetical imperatives tell us how to achieve our goals. They require us, on pain of irrationality, to do certain things, but only because such actions will get us what we want.

For instance, if my goal is to lose twenty pounds (as it often is), then reason requires me to forgo that pint of luscious coffee ice cream. If I want to get that Wall Street job, then reason requires that I line up a good summer internship. Reason demands that I look both ways at a busy intersection if I want to remain alive. These rational commands apply to me because of what I care about. I am irrational if I disregard them or act in a way that frustrates them.

But what if I decide that I don't care about weight loss or the benefits that weight loss can bring? What if I don't care about a Wall Street job? What if I want to die rather than live? In that case, I am no longer rationally required to pass on the ice cream, get the internship, or look both ways before crossing the street. These commands of reason are precarious. Their existence depends entirely on what I want. When my desires change, these rational requirements change or disappear.

Many people think that all rational requirements are like this—that they are all hypothetical imperatives. That's precisely what the first premise of the Amoralist's Challenge states: all of our reasons for action depend on what we care about.

Kant saw the implications of this argument very clearly, and knew that he had to challenge that first premise. In his jargon, what we need is to show how there can be such a thing as a **categorical imperative**. This is also a command of reason. But unlike hypothetical imperatives, categorical imperatives are rational requirements that do *not* depend on what we care about. They are requirements of reason that apply to everyone who possesses reason—i.e., everyone able to reflect on the wisdom of her actions, and able to use such reflections to guide her actions. Categorical imperatives command us to do things whether we want to or not, with the result that if we ignore or disobey them, we are acting contrary to reason—i.e., irrationally.

Kant thought that *all moral duties are categorical imperatives*. They apply to us just because we are rational beings. We must obey them even if we don't want to, and even if moral obedience gets us nothing that we care about.

One lesson Kant took from his consideration of the golden rule is that the basic rules of morality do not depend on our desires. If they did, then moral rules would fail to apply to everyone, since our desires can differ from person to person. This would make morality too variable, and make it possible for people to escape from their moral duty just by changing what they want. Kant thought that he was defending common sense when he claimed that morality is, in this sense, universal—that everyone who can reason must obey its commands.

If moral duties really are categorical imperatives, then we act rationally so long as we act morally, and we act irrationally if we disregard the demands of morality. Is that sort of view defensible? Can we really justify the claim that it is rational for everyone to act morally—even if we know that, for some people, moral conduct will only undermine their goals?

Kant thought he could do this. Consider his *Argument for the Irrationality of Immorality*:

1. If you are rational, then you are consistent.
2. If you are consistent, then you obey the principle of universalizability.
3. If you obey the principle of universalizability, then you act morally.
4. Therefore, if you are rational, then you act morally.
5. Therefore, if you act immorally, then you are irrational.

It does seem that rationality requires consistency, as the first premise asserts. And, as we have discussed, the principle of universalizability is a demand of consistency. So, while more could certainly be said about these first two premises, let us take them for granted here and focus on the third.

This is the claim that obedience to the principle of universalizability guarantees that our conduct is moral.

Its location in this argument tells us that the principle of universalizability is a crucial element in Kant's reply to the Amoralist's Challenge. He needs to successfully defend the principle in order to secure the claim that rational people are moral people, and immoral people are irrational. Can he do it?

Assessing the Principle of Universalizability

Unfortunately, the principle of universalizability fails as a general test for the morality of our actions. Look at premise 3 of Kant's Argument for the Irrationality of Immorality. It says that a maxim's universalizability is a guarantee of an action's rightness. That is false. We can act on universalizable maxims and still do wrong.

The principle of universalizability seems to be a very attractive way of pointing out how unfairness and inconsistency lead to immorality. So, for instance, when a thief robs a bank in order to gain riches, Kant can show why his action is immoral. If everyone acted on that maxim, there would be no money in the bank to steal, and the thief's goal could not be achieved. But what if the thief had robbed the bank in order to cripple it and put it out of business? If everyone acted that way, then the thief's goal *could* be achieved. So the principle of universalizability fails to condemn the robbery. And yet such an act is surely wrong.

Suppose that someone wants nothing more than to have a picture-perfect lawn. His basic maxim is to do whatever it takes to preserve the beauty of his grass. This would be strange, but that is no strike against his plans. The real thorn in this guy's side is the mail carrier, who always trots across the lovely yard on his way to delivering the mail. He has been warned, repeatedly, to use the sidewalk—to no avail. So the homeowner decides to take matters into his own hands. The next time the mailman tramples his lawn, the homeowner pulls out a gun and kills him.

That's clearly the wrong thing to do. But this man's maxim is universalizable. It tells him to take whatever steps are necessary (including killing) in order to preserve the beauty of his lawn. What if everyone were to act upon this maxim? Could the homeowner's goal be met? Undoubtedly. And so, if the principle of universalizability is true, this is no murder; it is justifiable killing. That has to cast serious doubt on the principle.

This unlikely example is just an instance of a general problem, one that we have seen before. Recall the case of the fanatic that arose in discussion

of the *What if everyone did that?* test. The goals of fanatics are ones that can often be met in a world in which everyone shares their aims. Fanatics (such as the lawn fanatic) need not make exceptions of themselves. The murderous aims of any number of groups could easily be achieved in a world in which everyone supported them. Thus fanatics can be consistent in the relevant sense: their guiding principles could be fulfilled if everyone else were to adopt them.

I think this shows that the principle of universalizability fails to give us an adequate test of fairness. For we can follow its advice while still singling out individuals or groups for discriminatory treatment. There can be consistent Nazis, after all. It doesn't follow that their policies are fair or morally acceptable.

Integrity

While utilitarians think of benevolence (the steady commitment to do good for others) as the central moral virtue, Kant touts integrity. Having integrity is living in harmony with the principles you believe in. It is the virtue of consistency. Integrity requires that you resist making an exception of yourself. It demands that you follow your principles even when doing so comes at a real cost. Kant is surely right that there is something admirable about integrity.

But integrity is not the only moral virtue, and it isn't even the most important one. The example of the fanatic reveals this. We may be absolutely dedicated to our principles, but if those principles are deeply flawed, it would be better, morally speaking, for us to have less integrity. That's surely what we want of the Nazi commandant or the dedicated terrorist. It would be better were they less principled people. We want them to be more flexible, and more open to the possibility that their guiding ideals are mistaken. When Huck Finn beats himself up for continuing to hide Jim, the escaped slave who accompanies him down the Mississippi, we applaud Huck's lack of integrity. A Huck with greater integrity is also one who would betray Jim's location. We want Huck to be less than fully conscientious, since that will mean Jim's freedom.

Integrity is worthy of our admiration only when it is tied to morally legitimate principles. The problem, as we have seen, is that people of integrity may still be doing wrong. Refusing to make an exception of myself is no guarantee that my principles are morally acceptable. It's not that consistency is worthless. But it fails as a general test for the morality of the principles we live by.

Kant on Absolute Moral Duties

Kant thought that certain sorts of actions are never permitted. Lying is one of them. In a much-discussed case, that of the inquiring murderer, Kant has us imagine a man bent on killing. This man knocks at your door and asks if you know the location of his intended victim. You do. Should you reveal it? If you do, your information is almost certainly going to lead to murder.

Kant thought you had two decent choices. Ideally, you'd just say nothing. That wouldn't help the murderer, and it wouldn't involve lying. But what if this isn't an option? What if you have to say something? In that case, you have to tell the truth—because you must never lie, under any circumstances.

I think that this is the wrong answer, and the interesting thing is that Kant's own theory does not require him to give it. Kant himself was so convinced of the absolute immorality of lying that he misapplied his own theory.

Kant never provided an argument for the claim that the moral rules that prohibit such things as lying and killing are absolute (i.e., never permissibly broken). The closest he came to supplying such an argument was in his belief that moral considerations are more important than anything else. In any conflict between moral duty and other demands—say, those of the law, self-interest, or tradition—morality wins.

Still, it doesn't follow that moral duties are absolute. For even if they always outweigh other kinds of considerations, moral duties might conflict *with other moral duties*. And if they do, they can't all be absolute. Some of them must give way to others.

And can't moral duties conflict with each other? It seems, for instance, that there is a duty to avoid hurting people's feelings, a duty to preserve national security, a duty not to start a panic, and a duty to protect innocent people from dangerous attackers. It also seems that fulfilling these duties will sometimes require us to lie, and that there is a moral duty not to do so. Perhaps none of these are really moral duties. Or perhaps, implausibly, we'd never need to lie in order to respect these duties. But much more likely is the thought that these are real duties, and that they really can conflict with another genuine duty—the duty not to lie.

But if that is so, then these duties cannot all be absolute. For if they were, then contradiction would ensue. Suppose, for instance, that the president is under an absolute duty to preserve national security. But if the only way to do that is to lie, and if lying is always wrong, then lying in this case would be both right (because it preserved national security) and wrong (because it

was a lie). And this is a contradiction. So if moral duties can conflict—and it certainly seems they can—then they cannot all be absolute.¹

This does not spell disaster for Kant. He does not need to defend the existence of absolute moral duties. His philosophy can, for instance, justify lying to the inquiring murderer. Kant's hatred of lying made him overlook a crucial element of his own view—namely, that the morality of action depends on one's maxim. He just assumed that anyone who lied would be operating with a maxim like this: tell a lie so as to gain some benefit. That maxim is not universalizable. In a world in which everyone did this, no one would trust the words of others, and people would be unable to secure any of the goals they were trying to achieve through lying.

But why think that Kant's maxim is the only one you could have in such a situation? A maxim is a principle that you give yourself. No one forces it on you. You decide what you are about to do, and why you are about to do it. For instance, suppose I used this maxim to justify lying to the inquiring murderer: say what you need to say in order to prevent the murder of an innocent person. I think that this maxim is universalizable. If everyone adopted this maxim, there would not be a general breakdown of trust and communication. People would still believe one another. The goal I am aiming for—to save an innocent person's life—could be achieved if this maxim were shared by everyone.

The maxim that tells me to lie to the inquiring murderer may be quite different from yours or anyone else's. And since, for Kant, the morality of an action depends on its maxim, we can't determine whether an act is right or wrong until we know its maxim. There is only way for Kant to absolutely ban a type of action. And that is to be sure in advance that every maxim that allows it fails to be universalizable. It is hard to see how we could ever know that.

This is all to the good, since it opens up the possibility that lying to the inquiring murderer is morally okay. Of course, if Kant is right, then we would have to have a universalizable maxim that permits this. But nothing Kant ever said should make us think that this is impossible. Contrary to Kant's personal view, we don't have to regard all (or perhaps any) moral duties as absolute.

Notes

1. For much more discussion of absolute duties, conflicts among them, and the nature of contradiction, please see chapter 15, pp. 211–212.

The Kantian Perspective: Autonomy and Respect

Is there anything wrong with slavery?

This probably sounds like an idiotic question. *Of course* slavery is wrong. So let me rephrase my question: Is there anything wrong, *in and of itself*, with enslaving other people? In practice, slavery has always created much more harm than good. But what if that were not the case? What if the members of a slave society—slaves as well as masters—were, on the whole, wealthier, better educated, healthier, and better satisfied with their lives than most members of a free society? And what if the abolition of slavery was sure to undercut these greater benefits? In those circumstances, would slavery still be wrong?

This thought experiment was put to readers by an important twentieth-century moral philosopher, Richard Hare. In his article “What Is Wrong with Slavery,”¹ Hare defended the utilitarian view that denied that anything is intrinsically wrong with slavery. Everything depends on the actual results of a slave system; in the imagined example, Hare had to admit that the slave society, since it created greater overall benefits, was the morally superior option. This despite the fact that Hare was, for all practical purposes, once a slave himself. As a British soldier in World War II, he was captured by Japanese forces and interned in a camp that effectively enslaved its inmates.

Hare emphasized that his views did not license any slave system as actually practiced. He presented the story as a way to show that there is nothing intrinsically wrong with slavery. The utilitarian says that the morality of slavery, like that of any other practice, depends entirely on its results. In the picture Hare paints, slavery can be morally acceptable.

Many will recoil at this verdict. They will feel that slavery can never be morally right, because it grossly violates people's autonomy. Slavery allows people to be treated as mere things—objects without any rights, of no intrinsic importance.

This is precisely the Kantian objection to slavery. Morality requires us always to treat human beings with the dignity they deserve. Slavery is inherently disrespectful. No one deserves such treatment. That is what explains why slavery is wrong.

Intuitively, this makes good sense. But it requires a bit of work to unpack it. We need to better understand why treating people as they deserve is so important, and what it means, specifically, to say that we deserve dignity and respect.

The Principle of Humanity

In the course of his work, Kant identified a number of different candidates for the role of ultimate moral principle. He thought that they each ended up requiring and forbidding precisely the same things, though most philosophers see important differences in these principles, and think that they sometimes issue different recommendations. While the principle of universalizability clearly emphasizes the moral importance of fairness, another of Kant's formulations directs our attention to the respect and dignity that serve as the basis of our moral treatment of others. This formulation is widely known as the **principle of humanity**:

Always treat a human being (yourself included) as an end, and never as a mere means.

To understand this principle, we need to get clear about three things: humanity, ends, and means.

When Kant spoke of *humanity*, he was not thinking of all members of the species *Homo sapiens*. Rather, he was (for reasons that will soon become clear) referring to all rational and autonomous beings. Perhaps there are aliens, or some other mammals, who are rational and autonomous. If so, then they count as human beings for purposes of Kant's principle.

Treating someone *as an end* is treating her with the respect she deserves. Treating someone *as a means* is dealing with her so that she helps you achieve one of your goals. This may be perfectly okay. I do this, for instance, when I hire a plumber to fix a broken water pipe in my kitchen.

In an innocent sense, I am using him—he is needed to get me what I want (a functioning sink, in this case). Yet if I greet him at the door, give him any assistance he asks for, and then pay him as he leaves, I am also treating him with respect, and so, in Kantian terms, I am also treating him as an end.

But what if, while the plumber is checking the leak, I remove a wrench from his toolkit and whack him over the head with it? He's out cold—excellent. I then snugly fit his head into the space where the pipe has corroded, thus temporarily stopping the leak. While he's unconscious, I rush off to the hardware store and buy a cheap bit of PVC pipe. The plumber wakes up just as I am returning from the store. I scold him for falling asleep on the job, and usher him out the door with a curt good riddance. Then I proceed to fix the leak myself, saving myself a hefty fee.

What has happened in this ridiculous scenario is that I've used the plumber literally as a thing, as a piece of pipe. He might as well have been an inanimate object. I failed to treat him in a way that recognized any of his distinctively human features. That's why I have treated him *as a mere means*.

While it often happens that people do treat one another both as an end and as a means, one can't treat people both as an end and as a *mere means*. Treating someone as an end implies a degree of respect that is absent when treating someone as a mere means.

Most of us think that there is something about humanity that lends us dignity and makes us worthy of respect. Most of us also think that human beings are worthy of greater respect than anything else in creation. Humans are more important than monkeys or sharks or daffodils or amoebas. Is this a defensible position, or is it just a self-interested prejudice?

Kant had an answer. He claimed that we are each rational and autonomous, and that these traits are what justify our special moral status. These two powers make us worthy of respect. Being rational, as we have seen, involves using our reason to tell us how to achieve our goals and to determine whether we can pursue them in a morally acceptable way. It takes a lot of brainpower to be able to formulate your goals, to imagine a world where everyone pursues them as you do, and then to ask about the consistency of your actions. Human beings are the only species on earth that can engage in such complex reasoning.

Being autonomous literally means being a self-legislator. Autonomous people are those who decide for themselves which principles are going to govern their life. You are an autonomous person. You possess the ultimate

responsibility for the choices you make, the goals you aim for, and the manner in which you pursue them. You are not a slave to your passions; you can resist temptation, check your animalistic urges, and decide for yourself whether to indulge them. You are not absolutely forced to act as you do, but are free to choose your own path.

Kant thought that our rationality and autonomy made each of us literally priceless. Despite the work of actuaries and juries in wrongful death suits, you cannot really put a dollar figure on a human life. Unlike mere objects, human beings are not replaceable one for another. The assumption that we are infinitely valuable explains our feelings of agonized loss at the death of a loved one. If we had to choose between the destruction of the most beautiful art object in the world and the killing of a human being, we should choose the former. No matter how valuable the object, the value of a human life exceeds it by an infinite amount.

The Importance of Rationality and Autonomy

Kant argues that rationality and autonomy support the dignity of each human being, and that everyone is owed a level of respect because of these traits. This makes excellent sense of a number of deeply held moral beliefs. Here are the most important of them.

1. It explains, in the first place, the immorality of a fanatic's actions. Such people don't regard human life as infinitely precious, but rather treat their despised opponents as mere obstacles to the achievement of their goals. The principle of humanity forbids such behavior, even when it is consistently undertaken, and thus allows us to address the most severe problem facing the principle of universalizability.

2. The importance of autonomy explains why slavery and rape are always immoral. Slavery treats the oppressed without regard for their own goals and hopes. Rape is treating another human being solely as a source of one's own gratification, as if the victim had no legitimate say in the matter. These are the most extreme examples of duress and coercion. They are immoral because of their complete denial of the victim's autonomy. As such, these crimes are perhaps the clearest cases of treating other people as mere means.

3. The principle of humanity easily explains our outrage at paternalism. To be paternalistic is to assume the rights and privileges of a parent—toward another adult. Paternalism has us limit the liberty of others, for their own good, against their will. It is treating autonomous individuals as children, as if we, and not they, were best suited to making the crucial

decisions of their lives. It is paternalistic, for instance, if a roommate sells your TV set because he is worried about your spending too much time watching *Seinfeld* reruns and too little time on your homework. Or imagine a classmate who thinks that your boyfriend is bad for you, and so writes him a nasty note and forges your signature, hoping that he'll break off your relationship. Anyone who has experienced paternalistic treatment knows how infuriating it can be. And the reason is simple: We are autonomous and rational, and the ability to create our own life plan entitles us to do so. We ought to be free to make a life for ourselves, even if, as is sometimes the case, we make a mess of things.

4. Our autonomy is what justifies the attitude of never abandoning hope in people. The chances that a very hard-hearted man will change his outlook may be very small, but the probability never reduces to zero. No matter how badly he was raised, or how badly he has lived his life, he is still autonomous, and so can always choose to better himself. It is usually naïve to expect such a transformation. Changing your character and habits is hardly easy. But the possibility of redemption is always there, and that is only because we are free to determine the principles that will guide our lives.

5. Many people believe in universal human rights. These are moral rights that protect every human being from certain kinds of treatment and entitle each of us to a minimum of respect, just because we are human. Kant can explain why we have such rights. We have them because of our rationality and autonomy. These two traits are the basis for living a meaningful life. If you doubt this, just imagine a life without them. It is a life fit for an insect, or a plant. What endows our life with preciousness is our ability to reason and choose for ourselves how we are going to live it. Every person is rational and autonomous to some degree, and every person needs these powers protected in order to have the sorts of experiences, engage in the kinds of activities, and support the sorts of relationships that make life worth living. Human rights protect these powers at a very fundamental level.

6. Our autonomy is what explains our practices of holding one another accountable for our deeds and misdeeds. Because we are not robots, but rather free and rational human beings, we are morally responsible for our choices and actions. We are fit for praise and blame, and that is because our conduct is up to us. We don't blame sharks or falcons for killing their prey; neither do we condemn a wilted orchid or a nasty-smelling ginkgo tree. Plants and animals deserve neither credit nor blame, and this is because their lives are not autonomous ones.

7. Relatedly, most people believe that punishment, rather than conditioning, is the appropriate response to serious wrongdoers. When dogs “misbehave,” we don’t try to reason with them. We try to condition them to change their behavior through a set of rewards and punishments. They don’t deserve to be punished when they break our rules, and that is because they lack the power to change their behavior by reasoning about it. By contrast, humans do sometimes deserve to be punished, precisely because they could have chosen to act well, but decided to act badly instead. People also deserve not to be manipulated into becoming obedient citizens. If we want criminals to behave differently, we must still respect their autonomy. The importance of autonomy explains why it is so objectionable to brain-wash people, or to drug or torture them into doing what we want.

The Good Will and Moral Worth

Kant’s insistence on the importance of rationality and autonomy led him to a view of intrinsic value that is very different from that of consequentialists. The structure of consequentialist thought is simple. Identify what is worth pursuing for its own sake; your moral duty is to maximize this value. Kant rejected this picture in every way.

Kant rejected the idea that happiness (or well-being in any form) is the ultimate value. Happiness has *no* value, he said, if it is experienced as a result of wrongdoing. (The enjoyment that a sadistic killer brings to his task does not add value to his crime, but only makes it worse.) And the same goes for other possible values. Wealth can be misused; so can power, and health, and understanding, and bravery. None of these is unconditionally valuable—none is valuable in every context. There is only one thing that is valuable, no matter what—only one thing whose presence in any situation is bound to add value to it. That one thing is the **good will**.

The good will has two parts. It is the ability to reliably know what your duty is, and a steady commitment to doing your duty for its own sake. The good will works in a familiar way: we see what we are morally required to do, and we do it for that very reason. No calculations of costs and benefits, no worries about what impression we might be making, what enemies we might be gaining, what riches might be in store for us. Once we understand where our duty lies, we do it straightaway.

Kant had some very interesting ideas about how the good will worked. Two of these ideas are especially important. Kant thought, first, that acting from the good will is the only way that actions can be truly praiseworthy.

(Kant referred to such actions as those that possessed **moral worth**.) He also thought that acting from such a motive is entirely an exercise of reason.

Consider the first point. Kant has us imagine two shopkeepers, each of whom does his duty by giving his customers the correct change. But the first does this only because he fears that if he were to cheat them, word would get out and he would lose business in the long run. He does his duty, but there is nothing morally worthy about his behavior.

The second store owner does the very same thing, but for completely different reasons. He treats his customers fairly because he thinks that cheating people is wrong, and he is committed to living up to the highest moral standards. This motivation earns the second shopkeeper the greatest praise. According to Kant, his actions and character display a worth that (like the value of humanity) is literally priceless. He is not for sale; he cannot be bought.

Kant's second point, about the importance of reason in motivating worthy conduct, is fairly complex. He thought that reason, operating alone and in the absence of any desires or emotions, could do double duty. It could reveal your moral duty, and it could motivate you to obey it.

To have a good will is, first of all, to know where your duty lies. Reason alone can tell you this. We can know what is morally required of us without the help of our feelings and emotions. When we determine whether a maxim is universalizable or think about whether a proposed action will respect the humanity in others, we don't need to want or feel anything at all. We just need to carefully follow the three-step test for a maxim's universalizability, or to reflect on the importance of autonomy. We can reason our way to moral knowledge. Indeed, neither our wants nor our emotions play any essential role in moral discovery. For Kant, we must be able to determine what is right and wrong by rational thinking alone, without the aid of desires or feelings.

That's because Kant saw these as unreliable moral guides. Compassion can lead you to wrongly help an escaping criminal; the courage of a terrorist can make his actions worse; anger can cloud impartial judgment. Our emotions often lead us astray, says Kant. They need to be guided by sound principles before we can trust them. Without such guidance, we *might* end up doing our duty, but that would be just a matter of luck.²

Further, and importantly, Kant thought that moral wisdom should be available to everyone, regardless of his or her emotional makeup. All of us are rational. We each have the power to reason well, even if we often fail to use this power as we should. But our emotions are not always under our

control, and they will differ from person to person. If a specific emotional makeup is needed to gain moral wisdom, then such wisdom might be out of reach for many of us. Kant thought that such a view is elitist and a denial of the fundamental equality of all human beings.

Knowing what you are required to do is one thing; actually doing it is another. Here Kant also downgraded desires and emotions in favor of reason. He denied the claim, made famous by Hume, that our motivations always depend on our desires. Hume thought that beliefs alone could never move us, and that we must want something before we will ever act. By contrast, Kant thought that we could do things even if we didn't want to do them, and even if we didn't think they'd get us anything we wanted. When acting from the good will, we are acting solely from an understanding of what is morally required of us, not from any desire or emotion. If our action is to have moral worth, then this understanding, all by itself, must be enough to motivate us.

Anticipating Freud by a hundred years, Kant argued that our motivations are hardly transparent. In fact, we can never be sure that we have *ever* acted from a good will. Still, even if we can't be sure that our actions have ever earned moral worth, we *can* know what standard we should aim for.

Kant went so far as to write that dutiful actions motivated by emotions or desires lack any moral worth. Those whose generous nature causes them to lend a helping hand are to receive no credit. Aid workers whose compassion or whose love of their work leads them to do what they do are not to be praised for their good deeds. But those who overcome a complete lack of interest and nonetheless offer help, not because they want to but just because it is their duty to do so, will receive full moral credit.

There are two ways to interpret Kant's message here. The first says that the presence of emotions is enough to rob an action of moral worth. The second is more charitable. It says that actions done solely from desire or emotion cannot possess moral worth, but that some cases of mixed motives—cases in which the good will moves us to act, though helped along by an emotional push—can yet have moral worth. Kant scholars are still conflicted as to which interpretation best captures his intentions.

Five Problems with the Principle of Humanity

Despite its many attractions, the principle of humanity, and the emphasis on rationality and autonomy that underlies it, are not trouble-free. There are five especially serious worries about the principle:

1. The notion of treating someone as an end is vague, and so the principle is difficult to apply.
2. The principle fails to give us good advice about how to determine what people deserve.
3. The principle assumes that we are genuinely autonomous, but that assumption may be false.
4. The principle assumes that the morality of our actions depends only on what we can autonomously control, but the existence of **moral luck** calls this into question.
5. The principle cannot explain why those who lack rationality and autonomy are deserving of respect.

Let's consider each of these problems in turn.

Vagueness

Unlike the three-step process used to apply the principle of universalizability, there is no straightforward test that tells us how to apply the principle of humanity. It tells us to treat humanity as an end—i.e., with the respect that people deserve. It's sometimes crystal clear whether the principle is being honored. No one doubts, for instance, that the principle is violated by treating a plumber as a piece of pipe or shooting a trespasser for trampling the lawn. But the vagueness of the notion of treating someone as an end often makes it difficult to know whether our actions are morally acceptable. Do we respect celebrities by telling the truth about their private lives—even when this is damaging to their reputations? Is it disrespectful to enemy soldiers to set land mines at our borders? Are we failing to give due respect to famine victims if we spend money on a new TV rather than giving money to an aid agency?

We can't know the answer to these questions without a better understanding of what it is to treat someone as an end. Without a more precise test of when we are respecting others and treating them as they deserve (i.e., as their rationality and autonomy demands), the principle of humanity fails to give us the guidance that we expect from an ultimate moral principle.

Determining Just Deserts

The second concern is about whether it is always appropriate to give people what they deserve. Kant certainly thought so. Recall his thinking, from the last chapter, about the prime importance of doing justice. Doing justice involves

giving people their just deserts—even if this is not going to benefit anyone. Sometimes this seems clearly right. A murderer ought to be punished, even if a governor's pardon will make more people happier. An employee ought to get paid for her work, even if her employer could do more good by giving her salary to charity. But there are also problems, as we'll see.

Kant has a partial reply to the problem of vagueness, mentioned earlier. He offers us a test for what *wrongdoers* deserve. That isn't the whole story, of course, since we also want to know how to apply the principle of humanity in cases where blame and punishment are not an issue. But even in contexts of condemnation, Kant's test—the famous **lex talionis**, or eye-for-an-eye principle—is fraught with difficulties. And so we are left with problems. In some cases, we don't know how to apply the principle of humanity, because it is unclear what treating a person as an end really amounts to. In other cases, it is clear—but also pretty clearly mistaken.

Lex talionis (the law of retaliation) tells us to treat criminals as they have treated their victims. Kant claimed that such punishment treats a criminal as an end—i.e., with the respect he deserves—because it treats him as a rational and autonomous person. Punishment is justified, for Kant, only if criminals are autonomous, and so able to freely choose their maxims. Those who are insane, for instance, are not fit for punishment. Punishment also presupposes that criminals are rational, in the sense of trying to act on principles that they can consistently intend everyone else to act on. A criminal's rationality permits us to turn his principles back on him, and do to him what he did to his victim. That is just what lex talionis requires.

Punishment that is administered as lex advises can be deeply satisfying. It can get criminals to see things from their victims' perspective, and so open their eyes to the true nature of the damage they have done. Further, punishment in line with lex seems perfectly just, since the criminal can't rightly complain of being mistreated. As Kant says, we would laugh at a criminal who protested against a punishment that harmed him exactly as he harmed his victim. Lastly, in the difficult matter of determining how to punish criminals, lex often gives us concrete, practical advice. What to do with a murderer, for instance? Kant counsels us to avoid the "serpent-windings" of utilitarianism and banish all thoughts of whether the death penalty is going to reduce the murder rate. A murderer deserves to die—lex says so. Therefore, morality requires his execution.

These attractions can explain lex's broad appeal. Despite the widespread enthusiasm, however, lex talionis is fatally flawed. Three reasons explain its failure.

1. First, *lex* cannot explain why criminals who intentionally hurt their victims should be punished more than those who accidentally cause the same harm. *Lex* tells us to set the punishment by reference to the suffering of the victim. But victims can suffer the same harm, whether the perpetrator has carefully planned to cause it or has caused it by accident. If I am recklessly practicing archery in my backyard and unintentionally skewer my neighbor, I deserve less punishment than a cold-blooded murderer. Or so we think. *Lex* does not allow for that, since the victims in both cases have suffered the same harm.

We could say that what criminals deserve is determined not only by the harm they have done, but also by how blameworthy they are in bringing it about. So a hired killer should be punished more than a reckless archer, because the murderer displays a kind of moral corruption that the archer lacks. This does give us the right answer—the callous killer *should* be punished more. But it comes at the cost of abandoning *lex*.

For we are no longer required to treat the criminal as he treated his victim. If an assassin deserves to be executed, then those who kill, but are less guilty than an assassin, should receive a lighter sentence than death. That undermines the letter and spirit of *lex talionis*, since these less-guilty killers will not be harmed just as they have harmed their victims. And it also removes one of the great virtues of *lex talionis*—that of offering precise guidance on how much criminals should be punished.

2. A second problem with *lex* is that it cannot tell us what many criminals deserve. This is most obvious in crimes that lack victims. Suppose an assassin attempts (but fails) to kill his victim, and the victim never discovers this. No harm, no foul? Suppose that someone leaves a bar well and truly drunk, and then manages to drive home without hurting anyone. Still, she deserves to be punished, but since there is no victim, *lex* offers no basis for punishment.

Other crimes may have victims, and yet *lex* offers no advice about their punishment. What to do with a hijacker or a counterfeiter? A kidnapper? Someone who transports stolen mattresses across state lines? The idea of treating these people just as they've treated their victims makes little sense.

3. Lastly, the guidance that *lex* provides, when it does prescribe a punishment, is sometimes deeply immoral. It's a sad truth: Any horror you can imagine people doing to one another has probably already been done. People have raped and tortured others, have burned whole families as they slept in their homes, have severed their limbs, tossed acid in their faces, and thrown handcuffed victims out of airplanes and helicopters. Does morality really require that we do such things to the criminals who

committed such deeds? Surely we don't want official torturers, rapists, and arsonists on the state payroll. Legal punishment is the state's business, and we insist that the state meet certain minimum moral standards. A state that rapes its rapists is failing, miserably.

These three problems show that *lex* cannot be the whole story about justice, because *lex* sometimes fails to give advice when it is needed, and sometimes gives bad advice. That means that when *lex* gets it right, it does so because its recommendations agree with those given by some more basic principle of justice. Homework: discover that basic principle.

In any event, most of us think that giving people the punishment they deserve sometimes has to take a backseat to other moral concerns. Our practice of allowing for parole, plea bargains, executive clemency, suspended sentences, and pardons attests to that. Each of these can be seen as an exercise in mercy—in treating people more kindly than they deserve. And mercy is a virtue. Kant's position requires that we never indulge in merciful treatment of criminals.

Suppose that maintaining a system of punishment required so much money that we had to drastically sacrifice funds for schooling, for health programs, and for national defense. In that case, perhaps we should punish criminals a bit less than they deserve, so as to save resources to meet these other social needs.

Suppose that punishing criminals as they deserve were to *increase* the crime rate rather than reduce it. Most would think this an excellent reason to lighten punishments.

Justice is very important. But these considerations should make us wonder whether Kant was right to think that justice must always be done, no matter its costs.

Are We Autonomous?

A third concern about the principle of humanity is that it is based largely on a questionable assumption—namely, that we are autonomous, that we are free to choose which principles to live by, and able to govern our actions by our choices. I do believe that I am autonomous. And so do you. We all do. But our confidence may be misplaced. *The Argument Against Autonomy* explains why:

1. Either our choices are necessitated or they are not.
2. If they are necessitated, then they are out of our control, and so we lack autonomy.

3. If they are not necessitated, then they are random, and so we lack autonomy.
4. Therefore, we lack autonomy.

Suppose our choices are necessitated—suppose, in other words, that we are determined to choose as we do. But how could that be? After all, don't you at this very moment have a choice about whether to put this book down or whether to continue reading? The choice is up to you. So long as your free choices are dictating your actions, you are autonomous.

But consider: Is anything influencing your choice? Of course. You choose to continue reading at least partly because you want to, because you believe you are able to, because you have no more appealing options, and because no one is forcing you to choose something else. *Given* all of these influences (and others, no doubt), it seems that you were bound to choose as you did.

True, it's not as if you are fated to keep reading *no matter what*. Rather, the idea is that you are destined to keep reading given your circumstances and your mind-set (your beliefs, desires, aims, etc.).

But the causes of your choice (your beliefs, desires, etc.) are also caused. These further causes don't spring up from nothing. You chose to keep reading partly because you wanted to. And you wanted to keep reading because (perhaps) you have been assigned this chapter and want to do well in the course you are taking. But your desire to do well in the course also has an explanation. It was caused by other desires and beliefs of yours, which, in turn, were caused by other factors, and so on, and so on. Ultimately, our choices can be traced to causes over which we lack control—causes such as our genetic inheritance, our parental upbringing, and a variety of social influences. If we choose as we do because of factors that ultimately are out of our control, then our choices are ultimately out of our control. And so, if our choices are necessitated to be what they are, then we are not autonomous.

Now suppose that our choices are not necessitated. Suppose that nothing really determined that you were going to choose to continue reading, for instance. You just chose to do so. If that were the case, wouldn't that show that you chose freely?

No. If nothing causes us to choose as we do, then our choices seem completely random. Randomness undermines control, and hence undercuts autonomy.

Suppose I'm walking down the hall and see someone thrust out her arm and hit a bystander. Did she choose to hit him? Yes. Why? No reason. No

cause. No explanation. It was just one of those things, completely out of the blue, unaccountable. But if that is really so—if nothing at all is causing her choice—then it seems that she isn't in command of her choices. They are something that happen to her, a passing fit of some sort, rather than something we can credit or blame her for. Her choices are out of her control.

Thus either way we go—whether our choices are necessitated or not—it seems that we lack autonomy. *If* that is so, the fundamental Kantian basis of our dignity, and the source of our duty to respect others, is undermined.

Of course many philosophers (and almost all non-philosophers) think that something is wrong with this argument. Its logic is watertight, so if there is an error, it must be in one of the three premises. Premise 1 is pretty clearly true. So the problem, if there is one, must lie in premise 2 or 3. Philosophers who think that we really do have autonomy have split on which premise to attack. Their work has been fruitful. (It has certainly multiplied—the issues of freedom and determinism nowadays form an entire subfield within philosophy.) So perhaps the pessimistic conclusion of this argument is false. But only a great deal more philosophy can show it so.

Moral Luck

Perhaps there is a flaw in the Argument Against Autonomy. I hope so. Let's indulge this hope for now and assume that we *are* genuinely autonomous. Even so, there are reasons to doubt that the morality of actions really depends on our autonomous choice.

Kant believes that we are rightly praised or blamed only for what we can control. That's why autonomy is so important. For autonomy *is* control—over our choices, and over our actions. Yet factors outside of our control apparently affect the morality of our conduct. If that is true, then autonomy may not play the central role in morality that Kant thinks it does.

The results of our actions are not fully within our control. And therefore they are morally irrelevant, from Kant's point of view. That is one of the main reasons that he so strongly opposes utilitarianism. And yet the results of our actions often *do* seem to make a moral difference.

Consider a good parent who, in a moment of extreme frustration, shakes her baby to jolt it out of a crying jag. Ordinarily, there is no lasting harm, the incident is forgotten, and we don't change our view of the mother's virtues. But babies sometimes die from such treatment. When they do, we judge the killer much more harshly than we do other parents—most of whom have shaken their babies at least once or twice, but have luckily done so without any permanent damage.

I sometimes find myself effectively driving on auto pilot. I've drifted over into the oncoming lane; I've failed to see a pedestrian at a crosswalk; I missed a passing car in my side mirror. In each case it was pure luck that my inattention didn't cause a (possibly fatal) accident. Many people are not so lucky. When their negligence results in someone's death, they are blamed far more than I have ever been. Yet they may be no worse a driver (or a person) than I am.

You probably know someone who is petty, vindictive, and coldly unsympathetic to the needs of others. A strict rule follower, someone who likes to pander to authority. Calculating and smart. This is the sort of person who can make office life pure hell. In decent circumstances, that is the worst he can do. Yet if he finds himself a citizen of an authoritarian regime, such a man can take the reins of a torture unit or a concentration camp and manage it with ruthless efficiency. It is simply a matter of luck that he is living in a peaceful society, rather than Germany in the 1940s, South Africa in the 1980s, or North Korea today. In different circumstances, we might well have charged him with complicity in torture or murder.

These are all examples of **moral luck**—cases in which the morality of an action or a decision depends on factors outside of our control. If Kant is right, moral luck cannot exist. And he may be right. But if he is, then we have to revise our moral views in each of these cases, and many others. A drag race down country lanes is fondly remembered—unless it leads to a paralyzing accident. The risky investment is harshly condemned if it forces bankruptcy, but celebrated if it establishes a family fortune. A revolutionary is a hero if his side wins, a despised traitor if it doesn't.

There is a tension at the heart of our moral thinking. We don't blame babies for any harm they cause. Adults who have been hypnotized or those who have been slipped an LSD tab are also immune from blame. Kant explained this perfectly: Such people lack autonomy. They aren't in full control of their actions. But if Kant is right, and control is essential for moral responsibility, then we must abandon all of the moral judgments recorded in the previous paragraphs. It isn't an easy choice to make.

The Scope of the Moral Community

The final concern has to do with the scope of the moral community. Kant's emphasis on rationality and autonomy forces us to draw the lines of this community very narrowly. We are in. Infants aren't. The severely mentally ill and mentally retarded are out. So too are all nonhuman animals and animal species, and all plants and ecosystems. They all lack rationality and

autonomy. By Kant's lights, they therefore have no intrinsic moral importance. We owe them no moral concern, and so, it seems, we can treat them any way we want.

We can express this worry in the *Argument Against Animals*:

1. If the principle of humanity is true, then animals have no rights.
2. If animals lack rights, then it is morally acceptable to torture them.
3. Therefore, if the principle of humanity is true, then it is morally acceptable to torture animals.
4. It isn't.
5. Therefore, the principle of humanity is false.

Though this argument focuses on animals, we could easily amend it to apply to infants, the severely mentally retarded, etc. Kant's views exclude all of them from the moral community. But since Kant himself focused only on the case of animals, let's follow his lead. We can discuss the other cases as needed.

Kant thought that it is wrong to torture or otherwise mistreat animals. So he accepts the fourth premise of the argument. He also accepts its first premise. He thought that rights require autonomy, that animals lack it, and that they therefore lack rights. As he saw it, the second premise is the one that has to go.

Kant offered two arguments for rejecting the second premise. Both of them fail.

He first claimed that harming animals will harden our hearts, and so make it likely that we will mistreat our fellow human beings. Since that really would be immoral, we must not harm animals.

Kant's predictions about how we might be led to harm our fellow human beings are quite shaky. Most of us are easily able to make distinctions in our treatment of members of different groups. Abusive bosses usually treat their superiors with respect. Ruthless prison guards can be loving parents. Doctors who are condescending to their patients and nurses are often quite decent to their fellow doctors. So mistreating one group needn't lead to mistreating others.

Further, if Kant is right, we humans really do possess infinitely greater moral importance than animals. Anyone who takes that message to heart would resist harming his fellow humans—even if he felt comfortable hurting animals.

Kant faces a problem even if his predictions are right. For the reasoning he is employing here is pure consequentialism. Don't torture animals, because that will have terrible results (it will lead to the mistreatment

of humans). But as we have seen, Kant bases his theory on the view that results are irrelevant to the morality of actions. So this reply will not do.

He has a second. I own a desk. It obviously isn't rational or autonomous. And yet no matter how much someone wanted to take a hammer to it, it would be wrong to do so. Not because the vandal would be wronging the desk, but because he would be wronging *me*. The desk has no rights. But I do. And these must be honored. And so, even though my three cats, for instance, have no rights, it would be immoral to hurt them, since in doing so, my rights (as their owner) would be violated.

There are two basic difficulties with this view. First, it offers no moral protections to wild animals. And second, domesticated animals will have no moral protection against their owners. If I decided to destroy my desk, just for the fun of it, I'd be doing nothing wrong. And since the Kantian view sees animals as morally on a par with my possessions, it can't explain why it would be wrong of me to destroy my animals simply because I wanted to.

That isn't the only bad news. Remember, this problem applies not only to animals but also to all human beings who lack rationality and autonomy. True, most of them (infants, the senile, the temporarily comatose, etc.) are loved by others. And so Kant might be able to claim that *our* rights (i.e., the rights of those who love such human beings) would be violated if anyone were to harm them. But what of the most piteous of humanity—the unloved, abandoned human beings who lack autonomy? Kant's theory gives them the same status as an unowned desk or animal. They are disposable and may be treated as we like. Kant thus excludes the most vulnerable among us from membership in the moral community.

Conclusion

Kant's ethical views are rich and suggestive. They are extremely important in their own right, but it can also be quite helpful to contrast them with the consequentialist outlook that is so popular in political and economic circles these days. As we have seen, Kant's opposition to consequentialism was deep and thorough. These are the main points of disagreement:

1. Kant denied that benevolence is the central moral virtue, and thought instead that justice and integrity occupied that role.
2. Kant regarded many of the basic moral rules as absolute, and so insisted that it was never acceptable to break them—even if breaking them led to better results.

3. Kant denied that the morality of actions could depend on results or other factors outside of one's control, and claimed instead that they depend solely on what we can be held responsible for—our maxims and our free actions.
4. In a related point, Kant rejected the exclusive emphasis on the future and an action's results in determining what is right and wrong, and instead made past actions, and their just deserts, a central basis for moral evaluation.
5. According to consequentialists, all it takes to be a member of the moral community is a minimal level of well-being; Kant thought instead that autonomy and rationality determined moral status.
6. Kant denied that happiness or well-being is always valuable in its own right, and instead believed that the good will—the steady commitment to do one's duty for its own sake—is the only thing that is valuable in all situations.

Many of the shortcomings of consequentialism are nicely handled by the Kantian theory. But consequentialists are pleased to return the favor: the Kantian theory isn't without its own problems, and many of those are neatly addressed by consequentialism. Let's now have a look at another important contender, the social contract theory, whose defenders hope to secure many of the benefits of these two ethical outlooks, while escaping the problems that confront them.

Notes

1. *Philosophy and Public Affairs* 8 (1979): 103–21.
2. This thought is perfectly illustrated by a catty remark quoted in a biography of Napoleon's sister, Pauline. She was notoriously pampered and unfaithful, an irresponsible spendthrift with a full sense of entitlement. She was sometimes capable of bravery and generosity. But this was unpredictable, and for the most part she behaved very badly. Later in life, a former acquaintance gave the following account:
"Although she was the most beautiful person one could imagine, she was also the most unreasonable.... talking inconsequentially, laughing at nothing and at everything, she contradicted the most serious people and put out her tongue at her sister-in-law when Josephine wasn't looking.... *she had no principles and was likely to do the right thing only by caprice.*" (My italics.)
As quoted in Flora Fraser, *Pauline Bonaparte: Venus of Empire* (Alfred A. Knopf, 2009), p. 25.

The Social Contract Tradition: The Theory and Its Attractions

The Lure of Proceduralism

One of the hardest and deepest problems for ethical inquiry is this: How to begin?

Should we assume, from the outset, that any plausible ethical theory will forbid rape and slavery and torture, and require compassion and kindness? If so, then we can use these assumptions as a litmus test. Any theory that contradicts these assumptions would be rejected. For instance, if killing or torturing an innocent person were ever in your self-interest, then ethical egoism would require you to perform such actions. It would therefore flunk our test. That would be enough to show that ethical egoism is false.

A question naturally arises: Can we justify our basic assumptions and defend the practice of rejecting any moral theory that doesn't agree with them? If a theory contradicts our basic assumptions about right and wrong, why does that undermine the theory, rather than our assumptions? Aren't we just begging the question against the ethical egoist, for instance, by assuming that self-interest could never justify torture or rape?

Defenders of this approach claim that there is no other alternative. We have to start moral thinking from somewhere—why not with those basic assumptions that almost everyone accepts, those that help support most of our other ethical claims? We can't prove that these assumptions are true. But that doesn't make them illegitimate.

Why not? Because the fundamental starting points in *every* area of thinking are beyond proof. It turns out, for instance, that the basic assumption behind all scientific inquiry—that there is a physical world outside of

our own mind—cannot be proven. All the evidence we might collect in order to prove it already assumes that such a world exists. Yet that doesn't render every scientific claim implausible or unjustified.

Many philosophers remain dissatisfied with this approach. They believe that we can justify our basic moral views, rather than simply take them for granted. We can do that by coming up with a *procedure* that will tell us the steps we must take to distinguish right from wrong. The correct moral views are those that emerge from the correct procedure. I call this sort of approach **proceduralism**.

Proceduralism tells us, for instance, that we should not begin moral inquiry by assuming, say, that slavery is wrong or that generosity is right. Make no moral assumptions at this stage. Instead, follow the correct procedure, and then see what the outcome is. Provided you've been careful, you'll land on the right answer to your moral question.

Of course, there is the matter of how to identify the correct procedure. Here's one way to do it: the correct procedure is the one that supports our deepest moral beliefs, and receives support in turn from these beliefs. But proceduralists are certainly going to reject *that* answer. They want to show us how to arrive at moral wisdom without first assuming the truth of our basic moral beliefs or principles. This is a very difficult challenge, as we'll see.

The golden rule is an instance of proceduralism. So is rule consequentialism. So is Kant's principle of universalizability. Each of these views tells us to follow certain steps in order to discover what is right and wrong. One of the main attractions of these views is that they do not take it for granted that slavery, for instance, is immoral; rather, they promise to *explain and justify* why slavery is immoral. They do this by showing, respectively, that (1) we wouldn't like it if we were enslaved, (2) no optimific social rule would permit slavery, or (3) no universalizable maxim would allow slavery.

If you have been reading these chapters in order, you know that there are significant problems with each of these proceduralist views. But that doesn't knock proceduralism out of contention, since there are many other proceduralist theories. Perhaps the most important of these is the social contract theory, nowadays known as **contractarianism**: the view that morality is based on a social contract.

The Background of the Social Contract Theory

Contractarianism originated as a political theory, and only later developed into a theory of morality. It tells us that laws are just if and only if they

reflect the terms of a social contract that free, equal, and rational people would accept as the basis of a cooperative life together. Its view of morality stems directly from that political ideal: *Actions are morally right if and only if they are permitted by rules that free, equal, and rational people would agree to live by, on the condition that others obey these rules as well.*

The theory's political origins can be traced back to the ancient Greeks. Early in the *Republic*, Plato's brothers tell Socrates that they find the social contract view both appealing and troubling. They challenge Socrates to tell them what is wrong with it. His answer takes up almost the whole of the book, a testament to the power of contractarianism.

Here is the story that Socrates heard. We are all by nature largely, or entirely, self-interested. What we want is power over others, physical security, plenty of money, and sensual pleasure. Our deepest goal: to lord it over everyone else. Who among us wouldn't want the power of the president or the wealth of Bill Gates—or, ideally, both?

This points to an obvious problem. Everyone wants to be at the top of the heap, and only a few can make it there. Further, no one wants to be a patsy, the person who gets stepped on as others climb the ladder of success. We each want to be Number One. But we know that the chances of making it are slim, and we want to avoid being trampled on as others claw their way to the top. So what do we do?

If we are truly rational, we will each agree to curb our self-interest and cooperate with one another. We'll do this *conditionally*—that is, on the condition that others do so as well. A complete free-for-all is going to make everyone miserable. If we all stop trying to get the better of each other, and instead agree to seek a little less for ourselves, then we'll all be better off.

That is what reason and morality require of us, according to the social contract theory. Starting with the assumptions that we each are largely motivated by self-interest, and that it is rational to be that way, contractarianism tells us that we each do best for ourselves by agreeing to limit the direct pursuit of what is best for us, and accept a bargain that gets us a pretty decent life. That everyone gets such a life means that we give up the chance of an absolutely fabulous life. But we also protect ourselves from a really terrible one, a life in which we are in the thick of a cutthroat competition, vulnerable to the attacks of everyone around us. That is a deal worth making. Here's why.

The Prisoner's Dilemma

Consider life's basic scenario: There is intense competition for scarce resources. We each want as much of those resources as we can get. Being

rational, we each try to get as much as we can, knowing that more for us means less for someone else. Things are going to get very bad, very quickly.

This is what happened to the thousands of fisherman who sought ever larger catches, resulting in the emptying of the Chesapeake Bay—for centuries one of the world's greatest fisheries. This is what happens when athletes take increasingly dangerous anabolic steroids, in a bid to gain a competitive edge and a lucrative championship. This is what happens when a politician starts a smear campaign and his opponent feels the need to ramp up the abuse in order to stand a fighting chance in the race. This is what always happens in turf battles over the spoils of an illegal drug trade.

These cases all share the same essential features. In each, there is mounting competition over a scarce resource, and many are trying their best to increase their share of it. That seems to be rational, and yet, if everyone stopped being so selfish, each person would be better off. There would still be fish in the Chesapeake, sustaining the communities surrounding it. Athletes would be safer, even if the world records in their fields were a bit less spectacular.

These sorts of situations, in which everyone would be better off by scaling back their pursuit of self-interest, are known as **prisoner's dilemmas**. The name comes from a scenario, introduced by economists, in which two thieves (call them Al and Bob) are caught and sent to separate detention cells. Being rational, Al and Bob previously made a deal with each other: If they get caught, they'll each keep silent, to thwart the police and protect themselves. Now that they have been captured, the police tell each one the same thing: "If you keep your promise to your partner by keeping quiet, and he rats you out, then he's off the hook, and you're looking at a six-year sentence. If *you* break your word and snitch on him, while he remains silent, you're home free, while he spends the next six years in jail. If you both keep quiet, you'll each get two years. But if you both confess, you'll each get four."

There is a diagram on the next page to help you keep track of the options. Each number represents years in jail. The first number in each pair is Al's prison sentence; the second is Bob's.

Suppose that both criminals know about the various outcomes, and that both have only one concern at this point: to minimize their jail time. If they are both rational, what are they going to do?

You might think that it's impossible to know the answer, since you don't know enough about Al or Bob, their bond with each other, their

		BOB	
		Remains Silent (Cooperation)	Confesses (Betrayal)
AL	Remains Silent (Cooperation)	2, 2	6, 0
	Confesses (Betrayal)	0, 6	4, 4

trustworthiness, etc., to make an informed guess. But really, there is no doubt that each is going to confess. They are going to break their promise to each other, landing themselves a four-year sentence apiece. That's a far cry from getting off scot free, and double the two years they'd get if they each kept quiet.

The important point is that remaining silent is the cooperative strategy. Silence here means keeping one's word, honoring the terms of the deal. Confession is a betrayal, breaking one's promise, abandoning a partner.

Al and Bob are going to betray each other. That's certain. They'll do this because they know the odds, because they are self-interested, and because they are rational.

Why will they confess? Because *no matter what his accomplice does, each criminal will be better off by confessing.*

Consider Al's choices. Suppose that

Bob remains silent. Then if Al confesses, Al is home free. If Al keeps his mouth shut, Al gets two years. So if Bob remains silent, Al should confess. That will minimize his jail time. That is what he most wants. So, if Al is rational, he will confess.

Now suppose that

Bob confesses. Then if Al confesses, Al gets four years in jail. Silence gets him six. So if Bob confesses, Al should confess, too.

Thus, either way, Al does best for himself by spilling the beans and breaking his promise to Bob. And of course Bob is reasoning in the same way. So they are both going to confess and end up with four years in jail.

The prisoner's dilemma isn't just some interesting thought experiment. It's real life. There are countless cases in which the rational pursuit

of self-interest will lead people to refuse to cooperate with one another, even though this leaves everyone much worse off.

The Chesapeake fishermen were in a prisoner's dilemma with one another. So are world-class athletes, once they discover that some of their competitors are taking performance-enhancing drugs. So are gang members who are gunning for their competition. A cooperative strategy would have saved them their livelihoods or their lives. So why don't they cooperate?

Cooperation and the State of Nature

The answer is simple: because it is so risky. The criminals in the prisoner's dilemma could cooperate. But that would mean taking a chance at a six-year sentence and betting everything on your partner's good faith. Unilaterally keeping silent, reducing your catch, refusing the use of steroids, forsaking violence—these are strategies for suckers. Those who adopt them may be virtuous, but they are the ones who will be left behind, rotting in jail, economically struggling, off the Olympic podium, or the victim of an enemy's gunshot. If enough people are willing to do what it takes to ensure that they get ahead, then you've either got to join in or be the sacrificial lamb.

Englishman Thomas Hobbes (1588–1679), the founder of modern contractarianism, was especially concerned with one sort of prisoner's dilemma. He invited the readers of his magnum opus, *Leviathan*, to imagine a situation in which there was no government, no central authority, no group with the exclusive power to enforce its will on others. He called this situation the **state of nature**. And he thought it was the worst place you could ever be.

In his terms, the state of nature is a “war of all against all, in which the life of man is solitary, poor, nasty, brutish and short.” People ruthlessly compete with one another for whatever goods are available, cooperation is a sham, and trust is nonexistent. Hobbes himself lived through a state of nature—the English Civil War—and thus had firsthand knowledge of its miseries. If you've ever read *The Lord of the Flies*, you have an idea of what Hobbes is talking about. As I write this, I can turn on my television and see pictures of states of nature from around the world—Somalia, Sudan, Afghanistan, the Gaza Strip. The scenes are terrible.

The Hobbesian state of nature is a prisoner's dilemma. By seeking to maximize self-interest, everyone is going to be worse off. In such dire

circumstances, everyone is competing to gain as much as he can, at the expense of others. With so much at stake, an all-out competition is bound to be terrible for almost everyone. No one is so smart or strong or well-connected as to be free from danger. No one is safe. There is no security and no peace.

There is an escape from the state of nature, and the exit strategy is the same for all prisoner's dilemmas. We need two things: beneficial rules that require cooperation and punish betrayal, and an enforcer who ensures that these rules are obeyed.

The rules are the terms of the social contract. They require us to give up the freedom to attack and to kill others, to cheat them and lie to them, to beat and threaten them and take from them whatever we can. In exchange for giving up these freedoms (and others), we gain the many advantages of cooperation. It is rational to give up some of your freedom, provided that you stand a good chance of getting something even more valuable in return. The peace and stability of a well-ordered society is worth it. That is the promise of the social contract.

But you need more than good rules of cooperation to escape from a prisoner's dilemma. You also need a way to make sure the rules are kept.

The state of nature comes to an end when people agree with one another to give up their unlimited freedoms and to cooperate on terms that are beneficial to all. The problem with agreements, though, is that they can be broken. And without a strong incentive to keep their promises, people in prisoner's dilemmas are going to break them. Just think of Al and Bob in our original example.

What's needed is a powerful person (or group) whose threats give everyone excellent reason to keep their word. The central power doesn't have to be a government—it could be a mob boss, who threatens Al and Bob with death if they were to break their silence. It could be the International Olympic Committee, with the power to suspend or disqualify athletes who test positive for illegal substances. But in the most general case, in which we are faced with anarchy and are trying to escape from utter lawlessness, what we need is a government to enforce basic rules of cooperation. Without a central government, the situation will spiral downhill into a battleground of competing factions and individuals, warlords and gang bosses, each vying for as much power and wealth as possible. A war of all against all isn't far behind.

Social contract theorists justify political authority—the power of a government to force its citizens to do things—by claiming that rational,

free, and equal people would agree to being governed. This agreement is easily explained. Who wouldn't prefer to live in a stable, peaceful society than in a lawless anarchy? Certainly, no government is perfect, and governments, like private citizens, can abuse their powers. Still, leaders earn the consent of the governed by making life better off than it would be in the state of nature. When they succeed in this, rational people will give them their allegiance.

The Advantages of Contractarianism

Contractarianism has many advantages. Here are some of the most important ones.

Morality Is Essentially a Social Phenomenon

For anyone all alone on a desert island or on a remote mountain retreat, there is no possibility of moral or immoral action. That is because moral rules are nothing other than special rules of cooperation, and when it comes to cooperation, it takes at least two to tango.

This explains why we have no self-regarding moral duties (duties that apply only to oneself). True, we can fail ourselves in a number of ways. We can become strung-out addicts and ruin our potential. We can make poor financial decisions that leave us bankrupt. We can make an awful career choice and spend our days in drudgery. In all of these cases we are highly imprudent. But when such actions have no bearing on others (they often do, of course), then contractarians will deny that there is anything immoral about them.

It Explains and Justifies the Content of the Basic Moral Rules

On the contractarian account, the moral rules are ones that are meant to govern social cooperation. When trying to figure out which standards are genuinely moral ones, contractarians ask us to imagine a group of free, equal, and rational people who are seeking terms of cooperation that each could reasonably accept. The rules they select to govern their lives together are the moral rules. They will almost certainly closely match the central moral rules we have long taken for granted.

John Rawls (1921–2002), the most famous twentieth-century social contract theorist, had a specific test for determining the rules that the ideal social contractors would support. In his *Theory of Justice* (1971), by most accounts the most important work of political philosophy written in the

twentieth century, Rawls has us envision contractors behind a **veil of ignorance**, an imaginary device that eliminates all knowledge of your distinctive traits. Those behind the veil know that they have certain basic human needs and wants, but they know nothing of their religious identity, their ethnicity, their social or economic status, their sex, or their moral character. The idea is to put everyone on equal footing, so that the choices they make are completely fair.

When placed behind a veil of ignorance, or in some other condition of equality and freedom, what social principles will rational people select? Almost certainly these will include rules that prohibit killing, rape, battery, theft, and fraud, and rules that require keeping one's word, returning what one owes, and being respectful of others. Contractarianism thus easily accounts for why the central moral rules are what they are—rational, self-interested people, free of coercion, would agree to obey them, so long as others are willing to obey them, too.

The rules of cooperation must be designed to benefit everyone, not just a few. Otherwise, only a few would rationally endorse them, while the rest would rationally ignore them. This allows the contractarian to explain why slavery and racial and sexual discrimination are so deeply immoral. Segregationist laws undermine the primary point of morality—to create fair terms of cooperation that could earn the backing of everyone. Even if, as is sometimes the case, oppressed people identify with the interests of their oppressors, and staunchly defend the system of discrimination, that does not make it right. The correct moral rules are those that free people would endorse for their *mutual* benefit—not for the benefit of one group over another.

It Offers a Method for Justifying Every Moral Rule

Contractarianism is a prime example of proceduralism. Contractarians have a method for seeing whether certain actions are right or wrong. They invite us to think about whether free, equal, and rational people would agree to live by rules that allow the action in question. Contractarians do not assume that the standard moral evils are bad; rather, they *show why* they are bad, by showing that rational contractors, getting together to select mutually beneficial rules of cooperation, would forbid such behavior.

Thus we are never at the point where we have to say to people, "Don't you *just see* that _____ (fill in the blank) is wrong?" Contractarians never have to make this move, which is excellent, since it can't possibly convince someone who doesn't agree with you. Rather, contractarians

offer a method for justifying every single moral rule. Of course disagreements will arise when we apply this method. There is plenty of room for argument about what free, equal, and rational people would accept. Still, if contractarianism is correct, then we will always have something to say when it comes to justifying even the most fundamental moral claims.

It Explains the Objectivity of Morality

Contractarianism offers an attractive picture of the status of morality. Moral rules, in this view, are objective. Anyone can be mistaken about what morality demands. Personal opinion isn't the final authority in ethics. Neither is the law or conventional wisdom—whole societies can be mistaken about what is right and wrong, because they may be mistaken about what rational contractors would include in their ideal social code.

Thus contractarians have an answer to a perennial challenge: If morality isn't a human creation, where did it come from? If contractarianism is correct, morality does not come from God. Nor does it come from human opinion. Rather, morality is the set of rules that would be agreed to by people who are very like us, only more rational and wholly free, and who are seeking to find terms of cooperation that will benefit each and every one of them.

Thus contractarians don't have to picture moral rules as eternally true. And they can deny that moral rules are just like the rules of logic, or of natural science—other areas where we acknowledge the existence of objective truths. The moral rules are the outcomes of rational choice, tailored to the specifics of human nature and the typical situations that humans find themselves in. This removes the mystery of objective morality. Even if God doesn't exist, there can still be objective values, so long as there are mutually beneficial rules that people would agree to if they were positioned as equals, fully rational and free.

It Explains Why It Is Sometimes Acceptable to Break the Moral Rules

Moral rules are designed for cooperative living. But when cooperation collapses, the entire point of morality disappears. When things become so bad that the state of nature approaches, or has been reached, then the ordinary moral rules lose their force.

One way to put this idea is to say that every moral rule has a built-in escape clause: Do not kill, cheat, intimidate, etc., *so long as others are obeying this rule as well*. When those around you are saying one thing and

doing another, and cannot be counted on to limit the pursuit of their self-interest, then you are freed of your ordinary moral obligations to them. The basis of morality is cooperation. And that requires trust. When that trust is gone, you are effectively in a state of nature. The moral rules don't apply there, because the basic requirement of moral life is not met: that everyone be willing to cooperate on fair terms that benefit all.

This explains why you aren't bound to keep promises made at gun-point, or to be the only taxpayer in a land of tax cheats. It explains why you don't have to wait patiently in line when many others are breaking out of it, or to obey a curfew or a handgun law if everyone else is violating it. When you can't rely on others, there is no point in making the sacrifices that cooperative living requires. There is no moral duty to play the sucker.

More Advantages: Morality and the Law

Contractarianism Justifies a Basic Moral Duty to Obey the Law

The social contract theory also has plausible things to say about why we are sometimes morally allowed to break the law. Before explaining this, though, consider why such cases are the exceptions, why we are usually, on contractarian grounds, morally required to *obey* the law.

The law enables us to escape from the state of nature, and so to gain all of the good things that come from a stable, peaceful society. And the benefits are real. We can securely purchase items on-line, safely walk the streets, trust the mail service, rely on legal contracts that we've entered, only because most people involved in these activities honor the rules that benefit everyone.

You have a role to play here. Your obedience to the law helps to support the institutions that make so many benefits possible. Those who break the law are undermining these highly beneficial institutions. They are also taking unfair advantage of the sacrifices made by their fellow citizens. That is immoral. Therefore, we all have a basic moral duty to obey the law.

The Contractarian Justification of Legal Punishment

Those who break the law are rightly punished. And the social contract theory has a natural explanation of this. Its account nicely blends some of the most attractive aspects of consequentialism and Kantianism, while solving a couple of hard problems in the philosophy of law along the way.

If punishment fails to deter crime, then the state cannot effectively serve its enforcer role, and so its ultimate justification is undercut. To

achieve its aim of keeping the peace, the state's threats for law-breaking have to be credible. To be credible, they must usually be carried out. And so, when people break the law, they need to be punished. That is the only way we can avoid falling into a state of nature.

Social contract theorists can also offer a Kantian rationale for legal punishment. When laws are good, they set out terms of fair play. When criminals break these rules, they take unfair advantage of their fellow citizens. They get all of the benefits of membership without shouldering the burdens that make the benefits possible. We right such wrongs by removing the criminal's ill-gotten gains. Punishment can do that. It can restore a level playing field. It can eliminate a criminal's unjust enrichment. It corrects the balance of benefits and burdens in society, and thereby sends a message that everyone is equal before the law.

Criminals have acted as if they could simply take what they wanted, without making the needed sacrifices that the rest of us have made. In doing so, criminals have placed themselves outside the protection of the law. They cannot complain if they are punished as a result.

Contractarianism Justifies the State's Role in Criminal Law

The contractarian justification of punishment neatly answers two related questions that have long perplexed philosophers: (1) why should the state, as opposed to private citizens, be the one who brings criminal charges and administers punishments? And (2) why should we have a criminal law in the first place? Why not just have the civil law (torts, contracts, etc.), and allow those who have been wronged to bring legal suits against those who have harmed them?

The social contract theory has answers. The state's ultimate purpose is to aid our escape from the state of nature. This gives it special authority to determine who is posing a threat to social stability. The actions that pose the greatest danger to society—treason, assassination, hijacking, attacking government agencies, etc.—are prohibited by the criminal law. That makes perfect sense, given the different purposes of the criminal and civil law. The function of the civil law is to repair personal harms and wrongs; the function of the criminal law is to preserve the state, and all of the advantages it provides. That is why a state representative (say, a district attorney or federal prosecutor), rather than a private citizen, initiates criminal charges. And that is why criminal punishment is a state affair, rather than a private matter to be settled between citizens.

Contractarianism and Civil Disobedience

Contractarianism has attractive things to say about why we have a duty to obey the law, and why criminals are rightly punished for violating just laws. But what if the laws are unjust, and fail to set out fair terms of cooperation? What if the ideal social code and a society's actual laws are miles apart? In that case, the social contract theorist can explain why *breaking* the law (under certain circumstances) can be morally acceptable.

Laws are morally justified when they mirror the rules that free and equal people would accept. And so, when laws themselves are grossly unfair, or when fair laws are applied in a discriminatory way, the primary point of having a society under law is undermined. Those who always get the short end of the stick are being constantly asked to sacrifice themselves for the sake of others. That isn't right, and the social contract theory explains why. The point of morality is to guarantee mutual benefit from fair terms of cooperation. Discriminatory policies and practices only frustrate this goal.

Governments must earn the allegiance of their citizens. They do so by making their lives much better off than they would be in a state of nature. That is why it is rational for people to agree to be governed. But when whole classes of people are enslaved or discriminated against, then the government has lost its moral authority over them. And that seriously weakens the moral duty to obey the government's laws.

Indeed, this means that it is sometimes appropriate to disobey the laws. The precise conditions are very difficult to state. But certain cases seem clear. Think, for instance, of the protests that Gandhi initiated against British colonial rule, those of Chinese democrats at Tiananmen Square, or those led by Martin Luther King, Jr., against segregationist laws in the American South. These cases of civil disobedience each involved illegal activities. But the laws being challenged were deeply immoral. And the illegal activities were morally powerful precisely because (1) the protestors took the law into their own hands only as a last resort, after all feasible negotiations had failed; (2) they acted openly, not secretly, and were willing to pay the price by going to jail and suffering the beatings of police; (3) they acted nonviolently, often in the face of physical violence; and (4) they were clearly motivated not by the promise of personal gain, but rather by the hope of furthering the cause of justice. In just about every respect, these classic cases of civil disobedience are morally very different from the work of ordinary criminals.

In acting as they did, the nonviolent protesters were showing a respect for the rule of law. This sounds paradoxical, but their message was clear. They tried to change the law because it was deeply unfair. They didn't want to replace it with anarchy, but rather with a system of real legal justice. They recognized the great value of a society under law—under just law. When there is no choice but to remain oppressed, resort to violence, or opt for peaceful and open resistance to injustice, the social contract theory favors the latter course. Isn't that the most admirable choice?



The Social Contract Tradition: Problems and Prospects

Why Be Moral?

Most moral theories have tried to show how, on their terms, it is rational to be moral, and irrational to be immoral. Social contract theorists are no exception.

The classic statement of this contractarian aim is given in Hobbes's *Leviathan*. Hobbes discusses the views of the calculating amoralist, whom he terms the Fool. The Fool admits that breaking his promises is unjust, but he doesn't care about whether his actions are just or not. He only cares about his own self-interest. When keeping his promises will do him good, he'll keep his promises. When it won't, he won't. He reasons that he'll sometimes do best for himself by being unjust. When that is so, it is rational to act immorally.

Hobbes faces the challenge directly. He agrees with the Fool's basic assumption: that self-interest is the fundamental reason for acting. In this line of thinking, you are rational to the extent that you increase your chances of becoming better off. Thus Hobbes has to show that acting morally is always likely to promote self-interest. That's not so easy.

The specific case Hobbes considers is one in which the Fool has made a deal with someone else, and the other person has already done what he's promised to do. Should the Fool keep his side of the bargain anyway?

Many will say no. Their advice is simple: take the money and run. That's certainly immoral. But people can get away with their immorality. It has happened many times. In such cases, people get what they want

without having to give up anything in return. It's a black eye for justice. But it seems the height of rationality.

Hobbes disagrees. He thinks there is a fundamental mistake in this reasoning. True, people sometimes get away with injustice. But that doesn't make it rational to act unjustly. Consider the fabled gang that couldn't shoot straight. They are a bunch of bumbler who do everything wrong, and yet through a series of lucky breaks, somehow manage to get away with their crime. They got what they wanted. But that doesn't mean that they acted rationally. Hobbes says that injustice is always like this. Even if the perpetrators sometimes get the goods, thinking that they can do it without getting caught and paying the price is always a bad bet.

Thus the question isn't whether immorality ever pays. Hobbes admits that it does. The question is whether acting unjustly ever increases the *likelihood* of personal gain. And Hobbes says that it doesn't. Though people can sometimes get away with injustice, the chances of doing so are never good. So injustice is never rational.

One way to defend Hobbes's conclusion is to raise the possibility of divine wrath. If you are immoral, then God is going to punish you, and no matter the worldly riches you're after, they simply aren't worth that misery. But Hobbes tosses this aside—the Fool he is considering is the same Fool who has said in his heart that there is no God. So the threat of divine punishment isn't going to work.

He imagines his Fool telling people of his plans to defraud them. That would indeed be stupid. But of course a clever and cunning criminal isn't going to do that. He'll keep his plans to himself, and try his best to hide his tracks so that his injustice is never discovered, or is pinned on an innocent person.

We can make the challenge hardest for Hobbes by considering a classic puzzle of rationality: the **free-rider problem**. This occurs when lots of folks are cooperating in a way that brings some common good. So long as enough people are chipping in, this benefit can be enjoyed by all—even those who refuse to help out. These are the free riders. They are getting a free ride by exploiting the efforts of others, without making any sacrifices themselves. This refusal to contribute to the support of the common good seems highly unfair. But it also seems highly rational.

Let's make this more concrete. Consider some shared public goods: keeping a park clean; maintaining a high level of decency in society; having democratic elections; eliminating a dreadful disease; having a secure national defense; having government agencies ensure the safety of our

food, medicines, and highways. These are genuine goods, and yet they can be achieved only if most people in a society are contributing to them. Such contributions involve sacrifice, sometimes small (making time to vote, not tossing a wrapper to the ground), and sometimes potentially large (serving in the armed forces, paying your tax bill).

In many of these cases, the contribution that any single individual makes is negligible. Democracy isn't going to crumble if I decide not to vote. The park won't be totally spoiled if I drop my gum wrapper on the grass. The government agencies will continue to do their work with or without my tax payment. Since that's so, I might reason as follows: The particular resource doesn't depend on my contribution; it will still be there whether I contribute or not. So, why make the sacrifice? I can gain the benefit of free elections or a secure national defense without taking the time to vote or serve in the military. I am not doing my fair share to support the common good. But I *am* getting something for nothing. How could that be irrational?

Of course, if you already care about being a moral person, and care about not exploiting your fellow citizens, then it will be rational for you to make the needed sacrifices to help support these common goods. You won't see voting, getting immunized, or paying taxes as a nasty burden, but rather as your contribution to causes that you wholeheartedly support. But suppose you don't feel this way, or you do, but are starting to wonder whether you are being a dupe. Hobbes doesn't assume that we are generous and public-minded. He is trying to show the person who is entirely self-interested why it would be a good bet, even for him, to do his fair share and to keep his promises. Can he do it?

There are two things we might say on Hobbes's behalf. First, we might argue, as he himself did, that the risks of doing wrong always outweigh its potential benefits. But this is simply mistaken. There are cases in which the chances of being found out are low, or the punishments are pretty mild and the benefits of injustice are really substantial. Since rationality, for Hobbes, is all about how much gain you can *reasonably expect* from an action, it can be highly rational in some cases to break your promises or let others make all the sacrifices.

What Hobbes could say is that it is never rational to behave unjustly in a *well-ordered* society. For him, such a society is one that offers reliable threats against breaking mutually beneficial rules. It won't be rational to cheat on your taxes, for instance, if the penalties for doing so are very severe and you are sure to get caught.¹

However, in our imperfect world, where enforcement isn't fool-proof, where punishments are sometimes pretty light, it can indeed be rational to take a chance at injustice. So long as we understand rationality as Hobbes does—as maximizing the chances of doing well for yourself—then he cannot show that injustice is always irrational.

But Hobbes might be able to show something nearly as comforting. Compare two claims:

1. No matter who you are, or what circumstances you find yourself in, it is always rational to act justly.
2. It is always rational to be a just person—the sort of person who values fairness, approves of just policies, tries to live an upright life, and becomes upset when learning of injustice.

Hobbes cannot defend claim 1. But he might be able to defend claim 2. And if he can, then he may also be able to show that

3. *For just people*, it is always rational to act justly.

And that would be a lot, even if it is a bit disappointing in leaving us without a surefire reply to the vicious person bent on doing evil.

Here is the best shot for Hobbes at defending claim 2. In the long run, just people are more likely to do better for themselves than unjust people. The upright person sleeps well at night. She has true friends. She has people she can share her delights with, be consoled by, confide her hopes and fears to. She isn't always looking over her shoulder, wondering whether the police are after her. She can review her life with satisfaction, knowing that she has earned the trust of others. The life of virtue has a lot going for it.

Compare such a life with that of an immoral criminal, one who is always worrying whether a vengeful victim or a turncoat conspirator is hard on his heels. It is a life of great insecurity. Those always on the lookout for the main chance, for the opportunity to cheat someone or exploit the system, are playing a losing game. They might win a few hands along the way, but there's not much in it as a long-term strategy. It isn't rational to be the cheating type.

If this is right, then it is rational to be a virtuous person. And if it is rational to be such a person, it is rational to try to stay that way. The best strategy for maintaining your good character is to keep on doing good things. For once we begin to give in to temptation, immorality can slowly chip away at virtue. It can be very hard to put on the brakes and prevent a slide into real corruption. Thus for good people, it is rational to continue

to act justly and to resist the call to stray from the path of virtue. That is precisely what claim 3 says.

That is the best case to be made for claims 2 and 3. And yet it also fails—or, rather, it isn't a complete success. It may be rational to be a moral person in societies that are well ordered, stable, and largely just. But when things turn bad, when corruption is rampant, when the powerful can get away with most things, then virtue can be a recipe for disaster. The ruthless have often taken advantage of good people's tendency to keep their word, to take the high road, to willingly sacrifice their own interests. Those who call for peace in the face of violence, who reveal the injustices of the dictator, who keep the rules as others are breaking them, are often the first to die.

Hobbes takes a hard-eyed view of human nature, and denies that justice is its own reward. If he is right about that, then justice pays only when it yields a longer life, a more secure existence, or a better chance at getting what you want. It sometimes will. But not always. When it doesn't, Hobbesian rationality requires that you make injustice your master.

The Role of Consent

Most of us believe that we have a moral duty to honor our commitments. And a contract is a commitment—it is a promise given in exchange for some expected benefit. A social contract differs from other contracts only in the extent of the duties it imposes and the benefits it creates. Since we are morally required to keep our promises, we have a duty to honor the terms of the social contract.

But have we actually promised to live up to any social contract? The Pilgrims did, when they paused before the shores of Massachusetts and together signed the Mayflower Compact in 1620. In ancient Athens, free men were brought to the public forum and directly asked to promise obedience to their city—or leave, without penalty. Naturalized citizens in the United States have long been required to pledge allegiance to the nation's laws. But relatively few adults nowadays have done any such thing. It seems, therefore, that we are not really parties to any such contract, and so are not bound to obey its terms.

Some philosophers have tried to finesse this point by arguing that we actually *have* agreed to obey the law. True, most of us haven't signed any contract or stated out loud that we are prepared to follow the law. Instead, we have offered our **tacit consent**, which is expressed through silence and a lack of opposition to the government. Tacit consent is a possibility: When I

ask my class whether it is okay to move on to a new topic, and no one replies, I reasonably take it that everyone agrees to proceed. Likewise, we might signal our consent to the government's laws by staying put, by not calling for its overthrow, by continuing to reap the benefits of civilized society.

But this is problematic. Some people loudly call for revolution; they are not tacitly consenting to obey the law. Others might be unable to voice their opposition freely, for fear of what the authorities will do. The fact that they remain where they are is no sign of their consent. It's simply evidence of a very practical decision. In many such cases, individual protest is unlikely to lead to better times, and is very likely to lead to imprisonment, torture, or an early death. How many political prisoners have learned this lesson the hard way? How many others have been deterred by their example?

The defender of tacit consent might reply that if people were rational, and saw the horrors of the state of nature, they would agree to being governed under the present terms, even if such terms are not ideal. And therefore, since people are, on the whole, pretty rational, we can assume that they have given at least their tacit consent to being governed by the existing laws.

Yet this is also mistaken reasoning. Suppose you make me a very generous offer, so generous that I'd be an idiot to turn it down. And yet that's just what I do. Perhaps I'm stubborn, short-sighted, or just plain stupid. You can't go on and insist that I uphold my end of the deal, since I never made one. Thus, even if it would be highly rational to consent to the social contract, that doesn't show that people have actually done so. And it is the *actual* consent of citizens, whether explicitly or tacitly given, that is supposed to be so important for justifying political authority, and imposing a moral duty to obey the law.

Thus it seems that many people have neither explicitly nor tacitly consented to the rules that govern their society. The reasoning here is summarized by the *Consent Argument*:

1. We have a duty to obey the law only if we have consented to do so.
2. Many have not given their consent to obey the law.
3. Therefore, many people do not have a duty to obey the law.

I have just tried to show that the second premise is true. Still, the argument works only if the first premise is also true. Its supporters claim that governments can be terribly abusive, and so their power can be justified only if it respects the will of its citizens. And that requires that citizens consent to being governed. An essential part of government is its power to force its people to do what it says. But as a general matter, it is wrong to

coerce people without their consent. Why should it be any different when the threat is coming from the government rather than a private citizen?

That is the strongest case for premise 1. Here is the strongest case against it.

Suppose you hate your country and reject its basic laws. You remain only because your government refuses to allow you to leave or because other, better countries won't let you in. You really, truly do not consent to your government's authority. Does that mean that you have no duty to obey its rules? That you are morally free to break laws against theft and slander and battery and murder, just because you renounce your membership in the social contract? That is the implication of premise 1. It is very hard to accept.

It is even harder to accept when we turn our focus away from questions of obedience to the law and ask, instead, about basic moral duties. It is *very* implausible to suppose that we are morally required to do only what we agree to do. We have a moral obligation to help the needy, even if we can't be bothered to think of their plight. I am duty-bound to offer support and gratitude to my loving parents, even if I never agreed to do so. Abusive bosses are morally required to be decent to their workers, even if they deny this and have made no promise to do so.

Contractarianism would be in deep trouble if it claimed that our moral duties applied only to those who agreed to accept them. *But it makes no such claim.* The social contract that fixes our basic moral duties is not one that any of us has *actually* consented to; rather, it is one that we each *would* agree to were we free and rational and seeking terms of mutually beneficial cooperation. So the fact that we have never signed a social contract or verbally announced our allegiance to one does not undermine the contractarian project.

Contractarianism does not require you to do whatever the existing laws and social customs tell you to do. Those standards are partly a product of ignorance, past deception and fraud, and imperfect political compromise. We are morally required to live up to the standards that free, rational people would accept as the terms of their cooperative living. It's safe to say that no existing set of laws perfectly lines up with those terms.

Thus contractarianism isn't a simple recipe to do whatever your society says. Rather, it provides a way to evaluate society's actual rules, by seeing how close (or how far) they are to the ideal social code that would be adopted if we were freer, more equal, and more rational than we are. If contractarianism is correct, this ideal social code is the moral law.

Disagreement Among the Contractors

If the social contract theory is correct, then the moral rules are those that free, equal, and rational people would agree to live by. But what happens if such people disagree with one another? For instance, what if these contractors can't reach a deal about the conditions under which a nation should go to war, or about the kind of aid we owe to the very poor? What happens then?

Rawls solved this problem by making every contractor a clone of every other. Behind the veil of ignorance, all of your distinguishing features go away. No one is any different from anyone else. And so there is no reason to expect any disagreement.

But Hobbes and other contractarians won't stand for this. They can't see why I should follow the rules of someone who is so completely unlike me—a person who is not only absolutely rational but also stripped of all knowledge of his social status, his friendships and family situation, his desires, interests, and hopes. Hobbes and his followers insist that the moral rules are those that we, *situated as we are*, would rationally agree to, provided of course that others would agree to live by them as well.

It's not easy to know how to solve this disagreement between contractarians. On the one hand, Rawls's view is likely to be fairer, since any information that could prejudice our choices is kept from us as we select rules to live by. But Hobbes also has a point, in that we want to make it rational, if we can, for everyone to live by the moral rules. Why should I live according to the rules set by some person who isn't at all like the real me? That's a pretty good question.

I'm sure that you've already figured out that I am not able to answer every good ethical question. This is another one I am going to leave for your consideration. Instead, let's return to our original problem: What should we say when the people choosing the social rules disagree with one another?

Perhaps Rawls is right, and there won't be any disagreement. But what if he's wrong? If contractors disagree, then the actions or policies they disagree about are morally neutral. They are neither required nor forbidden. That's because the moral rules are ones that *all* contractors would agree to. These rules are the product of a cooperative arrangement, one that expresses the support of all parties to the contract. If there are some matters that they can't agree on, then these are not covered by the moral rules.

This could be pretty bad. Or it might be just fine. It all depends on where the disagreement arises (if it ever does). If there are only small pockets of disagreement, regarding relatively trivial matters, then this is

hardly a problem. But what if contractors can't agree about war policy, about whether executions are just, about how to treat the poorest among us? Then this is really serious, since we do think that morality must weigh in on these issues.

So, how much disagreement will there be? There is no easy way to know. We can provide answers only after we know how to describe the contractors and their position of choice. Will they be clones of one another, situated behind the veil of ignorance? Or will they be aware of their different personalities and life situations? Will they be more or less equally situated, or are some going to have a lot more leverage than others? When we say that they are rational, do we have Kant's conception in mind? Or Hobbes's, according to which rationality amounts to reliably serving your self-interest? Or some other conception?

Answers to these questions will make a big difference in deciding on the specific moral rules that a social contract theory favors. They will also determine the amount of agreement we can expect from these contractors. But there is no shortcut to discovering these answers. To get them, contractarians must defend their own specific version of the theory against competing versions. That is a major undertaking. Until it is done, we cannot know just what the moral rules are, or how much contractual disagreement to expect.

The Scope of the Moral Community

The true measure of a man is how he treats someone who can do him absolutely no good.

—Samuel Johnson

Who has rights? Who deserves our respect? Utilitarians have their answer: anyone (or any animal) who can suffer harm. Kantians have their answer: anyone who is rational and autonomous. And contractarians have theirs: anyone whose interests are protected by the rules that contractors will agree upon.

When coming together to negotiate rules to live by, who will the contractors decide to protect? Different social contract theorists give different answers. The key to understanding them, however, lies in the idea that the contractors are, above all, *rational and self-interested*. It's easy to see why they should be rational. But why assume that they are self-interested?

There is a two-part answer. First, the preliminary part: being self-interested is not the same thing as being selfish. Being self-interested is having a strong concern for how well you are faring in life. Being selfish

is placing far too much importance on your own well-being relative to the interests of others. We don't have to assume that contractors are self-ish. But they definitely do care about how well off they will be under the agreed-on rules of cooperation.

Okay, fair enough. But the real question is whether we should think of the contractors as also being generous, benevolent, and self-sacrificing. And the answer that contractarians give is no. Why is that? Well, by assigning such virtues to the contractors, social contract theorists could be accused of stacking the deck in favor of moral principles that we already agree on. The attractions of proceduralism would be lost, since we would be starting our moral inquiry by making substantive assumptions about what is right and wrong. (Namely, that it is right to be generous, benevolent, and self-sacrificing, wrong to be the opposite.) And we would also be making it much more difficult to show each person why it is rational for him to obey the moral rules, since some people are not that virtuous. Some people would take a look at the rules that a bunch of awfully nice, sweet people agreed to, and roll their eyes in disgust. There's no way that they are going to live by *those* rules.

The assumption, then, is that everyone is to some degree self-interested and that it is rational to be that way.

Now back to our original question. If you are choosing from a position of equality, are free and highly rational, with a healthy dose of self-interest, who will you assign rights to? The answer is: other people like you. Contractors—those free, equal, and rational people engaged in the project of negotiating rules to live by—are the ones who will receive special treatment.

Contractors have a few defining features. First, they are both potential threats and potential benefactors. They can return our good deeds, but can also dish out some nasty treatment in response to being harmed. Second, they are, fundamentally, our equals. They have roughly the same powers as we do, including powers to help us and to hurt us. And third, *we must be unable to gain anything from them without their consent*. In other words, we must enter agreements with them in order to obtain benefits from them, and protection against them.

At first blush, it seems strange that membership in the moral community should be limited to contractors. But it makes good sense, once we recall the Hobbesian view of rationality. According to Hobbes, *sacrifice requires compensation*. When we escape from the state of nature, for instance, we give up a great many freedoms. But it is worth it, because of the promise of a better life, the sort that comes from being able to trust and rely on others.

Now here is the crucial question: If we can get what we really want, without having to sacrifice anything, then why should we make the

sacrifice? If Hobbes is right, there is no answer to this question. There is no reason to make such a sacrifice.

We can get what we want from trees, from animals, from the very weak, without having to give up anything in return. And so we have no reason to treat such beings with respect. They don't meet the conditions of being a contractor. We *could* decide to be nice to them—we could choose, for instance, to let a chicken live, rather than kill it and eat it, or let a lab rat go free, rather than submit it to painful experiments. But that is charity. It isn't duty. We have no duties to such vulnerable beings, because our duties are limited to contractors—those we have reason to respect, because they can harm us if we don't and can benefit us if we do. We can't just take what we want from contractors. We need to give up something in exchange for their beneficial cooperation.

In short, you get moral status, have rights, and are owed respect just in case your cooperation is needed in order for others to benefit. You have to be a member of the social contract in order to have genuine moral importance. Plants aren't members, since we can use them as we like, without giving them anything in return. Ditto for animals, and the very weak among us. That is why membership in the moral community is limited to contractors. They are the only ones it is rational to enter a social contract with. For all other beings, we'd be giving something up (freedom to exploit them), without gaining anything in return. According to Hobbes, that is irrational.

Thus contractarianism offers no secure protection for the truly vulnerable. We could, if we like, ensure their protection, but reason does not require us to do so, since we can get what we want from them without having to sacrifice anything in return. We are to them as a much more powerful and indestructible alien race would be to us. Such creatures, if they felt like it, might be nice to us. But since there is nothing in it for them to treat us this way, neither reason nor morality requires that they offer us their respect. They'd be perfectly within their rights to snack on us day and night.

This is because morality, in the social contract view, comes from rules of beneficial cooperation that rational parties would agree to. Rational, all-powerful aliens wouldn't agree to limit their liberty in exchange for anything we might do for them, since we can't hurt them, and they can get whatever they want from us without giving anything up on their end.

If the social contract theory is correct, then animals, ecosystems, infants and the severely retarded are in the same position as we readers are in the fictional example I've just presented. It's very bad news for us if such a race ever pays the earth a visit. In the meantime, it's very bad news indeed for all the vulnerable among us.

Conclusion

Contractarianism starts with a very promising idea: Morality is essentially a social matter, and is made up of the rules that we would accept if we were free, equal, and fully rational. The heart of the theory is an ideal social code that serves as the true standard for what is right and wrong.

This theory has a lot going for it, as we've seen. It offers us a procedure for evaluating moral claims, and so offers the promise of being able to justify even our most basic moral views. It has an interesting explanation of the objectivity of morality. It can explain why we are often bound to obey the law, and why we may sometimes break it. It has an appealing account of why we should punish criminals. It can explain why it is usually (though not always) rational to behave morally.

That's a very impressive list. Like all other popular moral theories, however, it is not perfect. There are those who are disappointed with the theory because it leaves open the possibility that immoral actions are rational. Whether this is a serious problem, however, is still up in the air, since it is very hard to know whether it is always rational to do your duty. Still, as we have just seen, the theory does have at least one truly worrying aspect. It denies that the neediest are members of our moral community, and so opens the way to their exploitation.

I want to remind you that every moral theory we consider in this book is still a work in progress. Very smart philosophers continue to develop these theories, building on their strengths and trying to shore up their weaknesses. It's no different here, when it comes to assessing the social contract theory. I haven't offered the last word on its success. I've only offered the beginnings of the full picture. Those who find its vision of morality elegant and compelling are invited to think more about how it might be supported, while fans of its competitors must do the same for their favored theory.

Let's now have a look at more of the competition.

Notes

1. It is interesting that in Hobbes's line of thinking, many criminal organizations will count as well-ordered societies. Think, for instance, of how badly police informants fare in the Mafia society or drug culture depicted in *The Sopranos* or *The Wire*. Going over to the feds or the cops is almost always a very bad bet, and that's because the punishments for doing so are as bad as they can be, and pretty reliably enforced.



Ethical Pluralism and Absolute Moral Rules

The Structure of Moral Theories

All of the moral theories we have considered so far have one thing in common: they are examples of **ethical monism**. Monistic theories argue that there is one supreme rule that serves as the basis of all morality.

A supreme moral rule has two defining features. First, it is **absolute**. That means that we are *never* permitted to break it. If you violate an absolute rule, you have automatically acted immorally. Second, this moral rule is **fundamental**. There are no deeper, more basic moral rules that justify the supreme rule.

Utilitarians, for instance, are ethical monists. They insist on a single, ultimate moral rule: maximize happiness. This rule is absolute. Every act that fails to maximize happiness is wrong. And the rule is fundamental. If you ask utilitarians to justify this requirement, they will deny that any other moral rule can do so. Moral questioning, like all other lines of investigation, must stop somewhere. Fundamental moral rules mark that stopping point.

Contrast this with the moral rule that requires surgeons to use anesthesia on patients, or the rule that forbids children from teasing their unpopular classmates. These rules are not fundamental ones. Their justification comes from their link with a more general and basic rule, such as one that requires us to avoid inflicting unnecessary suffering or to show respect to our fellow human beings.

Almost all of the classic moral theories are monistic. They each defend a single absolute, fundamental moral rule (maximize self-interest, maximize

happiness, do what God commands, act only on universalizable maxims, etc.). They then use this fundamental rule as the supreme test of morality.

The attractions of a monistic theory are clear.¹ We naturally seek unification in our thinking, and monistic theories provide this. They can impose order on our moral thinking, and organize all moral principles by reference to a supreme moral rule. But we have reviewed the most important versions of ethical monism and have found serious problems for all of them. What to do?

There are three options to choose from: (1) We could try to discover a new version of monism that puts forward a supreme moral rule that we haven't seen before; (2) we could stick with one of the theories we have already seen, and try to perfect it by defending it against objections; or (3) we could abandon the monistic assumption that has driven so much moral philosophy. This third option is the path of **ethical pluralism**. We will explore its central claims in this chapter and the next.

Ethical pluralism is a family of views that holds that there is a plurality of fundamental moral rules. Thus pluralists deny that we can systematize ethics under a single rule. Some pluralists believe there are two basic moral rules; others believe that there are three or more. There are many different versions of pluralism, each one distinguishing itself by the different rules it considers fundamental.

Luckily, we can spare ourselves the effort of examining every variation (too many to count, in any event), and divide pluralistic theories into two camps. The first is that of the absolutists, who think that it is always wrong to violate the fundamental moral rules. Other pluralists reject this view. They think that it is sometimes morally acceptable to break a fundamental moral rule. Let's spend some time with the absolutists first, and then turn, in the next chapter, to their opponents.

Is Torture Always Immoral?

You might reasonably think that if there are any absolute moral rules, then a ban on torture would be among them. Let's consider a case to help sort things out.

On the evening of March 1, 2003, CIA officers deployed just outside of the Pakistani capital of Islamabad received a text message: "I am with K.S.M." The initials were those of Khalid Sheikh Mohammed, mastermind of the September 11 attacks on the World Trade Center. After waiting a few hours, so as not to compromise the identity of their informant, CIA

officers stormed the house and apprehended Sheikh Mohammed. He was soon flown to Afghanistan and then to Poland, where he was detained in a "black site" facility—a detention center that did not officially exist and that was run without legal oversight.

Over the next two weeks, Sheikh Mohammed was sometimes prevented from sleeping for extended periods, was slapped, was subjected to frigid temperatures, and was "waterboarded" more than one hundred times. Waterboarding is especially terrifying. It involves strapping a prisoner to a board and repeatedly pouring water over his mouth and nose to simulate the experience of drowning. As water fills his lungs, the prisoner almost immediately suffers a gag reflex and is overwhelmed by a feeling that he is about to die.

Waterboarding rarely results in death. It is always extremely painful, and it does sometimes kill, though it more frequently results in other sorts of harm (lasting psychological damage, lung or brain damage, and sometimes broken bones caused by struggling against the physical restraints).

In Sheikh Mohammed's case, the purpose of such treatment was clear—to extract useful information that would lead to the capture of other terrorists and the prevention of their murderous plans. Despite serious reservations by a number of CIA and FBI personnel, the most senior U.S. officials regarded such treatment as morally acceptable. Their argument was straightforward: Though waterboarding isn't exactly a noble undertaking, it is the lesser of two evils. If faced with a choice between torturing a terrorist, and saving the lives of hundreds of innocent people, it is clear which option is best.

It is interesting, though, to note how the public defense of such interrogations proceeded. Administration officials repeatedly denied that waterboarding is a form of torture. For they appeared to embrace the absolutist idea that torture is always immoral. That is why they argued so strenuously that waterboarding did not qualify as torture—had they admitted it, they would have felt bound to say that their actions were wrong.

But is there really an absolute moral ban on torture? Despite widespread support for this view, it might not stand up to scrutiny. Perhaps we should accept that waterboarding is torture, but insist, forthrightly, that torture is sometimes morally acceptable—especially when many innocent lives are at stake.

Politically, such a move carries great risks. It is the rare administration that is willing to own up to its use of torture, as there are very serious consequences of doing so. But we can leave political concerns aside and ask,

directly, whether torture is always morally wrong. When vividly imagining bouts of torture, it can be hard to resist the thought that such actions, of necessity, are immoral. And yet, as we shall see, it can be very difficult to justify this view.

Preventing Catastrophes

Those who condemn all torture must be prepared to answer the perennial challenge of the ticking bomb terrorist, the one who knows the location of a bomb powerful enough to kill thousands of innocent people. Suppose that we were able to capture such a terrorist and he refused to reveal the bomb's location. Wouldn't it be acceptable to torture him?

We should consider this challenge in its best light, so suppose that any torture we inflict will not be lethal and must first be authorized by at least two independent government officials. Further, it will be administered only after the prisoner has been given a chance to cooperate, and if he does, no torture will occur. Suppose all of these conditions have been met, but the prisoner refuses to cooperate. We then begin to torture him by means of an extremely painful technique that will cause no lasting physical damage—perhaps a series of sanitized needles inserted very deeply under the prisoner's fingernails. Given the stakes, isn't this sort of action justified?

Those who say yes would do so for an obvious reason: Torture in such a case might prevent a catastrophe. The terrible pain that we are proposing to inflict is done only because it is needed to stop a genuine disaster. The critic of absolutism will try to generalize from this case, and can offer the following *Argument from Disaster Prevention*:

1. If there are any absolute moral rules, then we are *never* permitted to break them.
2. Every moral rule may be permissibly broken, if doing so is necessary to prevent a catastrophe.
3. Therefore, there are no absolute moral rules.

This is a very powerful argument. Its first premise is certainly true, since absolute moral rules are, by definition, those that we are never allowed to break. The second premise is also very plausible—so plausible, in fact, that those who deny it face an extremely heavy burden of argument. What can possibly be so important about obedience to a moral rule that would allow us to sacrifice thousands, perhaps millions, for the sake

of keeping our hands clean and obeying it? That is a very hard question to answer. If there is no satisfactory reply, the conclusion must be true.

Notice that this argument makes no specific mention of torture. It is designed to be perfectly general, to undercut all absolute moral rules. But even if you think that the ban on torture should allow exceptions, perhaps you are unconvinced about other cases. Can it ever be right, for instance, to rape someone? Or to deliberately kill an innocent person against his wishes?

Critics of absolutism will say that, yes, such actions may indeed be morally right, though they will also admit that any such cases are probably extremely rare. For instance, if someone credibly threatens to detonate a nuclear bomb in a densely populated city, and agrees to defuse it on only one condition—that we authorize the killing of an innocent person—then perhaps, morally speaking, we must agree to his terms. We should never celebrate having to torture, rape, or kill. But if doing such actions will prevent a horrible catastrophe, then critics of absolutism insist that we may be permitted, and perhaps even required, to do such things.

Now there is a ready reply to this criticism. We could simply add a special clause to all of our moral rules, which requires that we obey them *except in cases of emergency*. But what counts as an emergency? That ten thousand lives are at stake—this is clear. But what if ten are in jeopardy? Five? Two, or one? What if the danger is not life-threatening, but we know that obedience to a rule will cost someone a limb, or the job that she has worked ten years to obtain?

The notion of a moral emergency is very vague, and unlikely to be of much use in helping us to know when exceptions to rules are called for. Once we allow any exceptions at all to moral rules, we are engaged in the very tricky business of trying to strike a decent balance between competing claims. It may be best, *if we are forced this far*, simply to abandon absolutism altogether and to accept that the best version of ethical pluralism is one that regards its fundamental rules as non-absolute.

The Doctrine of Double Effect

Suppose that you are an emergency room physician and are faced with six patients who have just been rushed to your hospital. They were all at a dinner party and made the mistake of eating some innocent-looking mushrooms that turned out to be poisonous. Each of the patients is dying. After quickly running some tests, you discover that saving one of the patients requires giving him all of the available antidote. You also discover that you

can save the other five patients by giving them each a fifth of the serum. Whom should you save?

Assuming that each of these six patients is a good person, with otherwise decent life prospects, it seems clear that you should save the five. It's not that the very needy patient deserves to die; rather, it is better that only one dies, rather than five.

We reach similar verdicts in other cases. Imagine that a runaway trolley is headed for five innocent people and that you can pull a lever that switches the trolley to a side track, where (you guessed it) one unfortunate person is trapped. Shouldn't you pull that lever?

Suppose that an important military operation can be conducted risking only one soldier's life, rather than five. If the one soldier is just as likely to complete the task, it seems clear that the right thing to do is to spare the five and send him out alone.

Reasoning in these cases, as in so many others, seems to push us directly to consequentialism.² The mandate here (and in the Argument from Disaster Prevention) seems clear: minimize harm.

Yet we would also minimize harm if we were secretly to abduct a small number of healthy people, anesthetize them, and cut them up to distribute their vital organs to those who would otherwise die from organ failure. We could minimize misery by "culling" the population of those whose lives are wretchedly unhappy, with little prospect of improvement—even if they didn't want to die. We could dramatically reduce terrorism if we adopted a policy of reliably executing a terrorist's child or spouse in response. But these ways of minimizing harm are deeply offensive.

Consequentialist reasoning, so powerful in the Argument from Disaster Prevention and in the three initial cases we've considered here, seems to have led us astray. But if it is so plausible in the earlier cases, perhaps we should overcome our queasiness when it comes to the examples in the previous paragraph.

Absolutists reject that thought. Indeed, many people are tempted by absolutism precisely because they see it as the only way to resist consequentialist reasoning. But to make good on this hope, absolutists need to show how they can make sense of these various cases, without relying on hidden consequentialist assumptions. The absolutist must be able to distinguish the antidote, trolley, and military stories, on the one hand, from the murderous cases just described.

Absolutists have often relied on the **doctrine of double effect (DDE)** to do this. The doctrine refers to two relevant effects that actions can have:

those that we intend to bring about, and those that we foresee but do not aim for. This principle says that

Provided that your goal is worthwhile, you are sometimes permitted to act in ways that foreseeably cause certain harms, though you must never intend to cause those harms.

The DDE does not say that it is always wrong to intentionally harm others. It allows, for instance, that harmful punishment is sometimes acceptable. The DDE simply tells us that *some* harms may never be aimed for, even though such harms may be permitted as side effects, or “collateral damage,” of one’s actions. Which harms are these? The ones prohibited by absolute moral rules.

The doctrine of double effect, if true, would have two very important implications: (1) it would provide a reply to the Argument from Disaster Prevention, and (2) it would refute utilitarianism and all forms of act consequentialism.

A Reply to the Argument from Disaster Prevention

The Argument from Disaster Prevention advises us to impose certain terrible harms in order to prevent even more of them from occurring. Those who endorse this argument would have us set out to kill innocent people, for instance, if doing so would lessen the overall number of innocent deaths.

Against this, the DDE requires that we not aim at certain harms, even if that is the only way to prevent disaster. After all, the absolutist insists that certain acts must never be done, *whatever the consequences*. The DDE thus offers absolutists a way to distinguish cases in which it is and isn’t morally acceptable to minimize harm. You may minimize harm, provided that you do so without intending to do the things that are absolutely forbidden to do. Thus the ends, even if they are valuable, do not always justify the means, if the means require that we intend to do an action that is absolutely forbidden.

How the DDE Threatens Act Consequentialism

If the DDE is correct, then even if two actions predictably have the same bad results, one of these actions might be right and the other wrong. And that is something that utilitarians and other act consequentialists cannot allow. They determine the morality of an action based solely on its results. So any two acts with identical results must be morally equivalent. By contrast, the DDE determines the morality of actions based partly on what is

going on in the person's mind. If the DDE is correct, then some intentional harms may be immoral, even though causing the same harms while aiming at other goals may be permitted.

The DDE can be used by absolutists in the following way. Suppose, as many absolutists claim, that it is always wrong to deliberately kill an innocent person. Then it is always immoral to intend to kill a healthy patient to save others, to aim to kill an innocent child to stave off terrorism, and to seek to eliminate the deeply miserable among us. Since each of these actions involves the intention to kill innocent people, it is automatically wrong.

But by sending a soldier on a dangerous reconnaissance mission knowing that he is likely not to return, we are aiming at some morally acceptable military goal rather than the soldier's death. In providing the antidote serum to five patients, we know that the neediest patient will die, but we would be delighted if he survived; we don't intend to kill him. In shifting the trolley to a side track, we are trying to save five lives, even though we foresee the cost of doing so. Still, if the innocent on the spur were able to find an escape, we wouldn't at that point set out to kill him! Our intention in this case is to save the five, not to kill the one. These last cases all involve morally acceptable actions; even though we foresee, in each case, that someone will probably die, we are aiming to do good, and do not intend to cause harm.

The DDE thus allows us to morally distinguish cases that yield identical results (in these cases, five lives saved, one lost), and does so by showing how we can respect an absolute ban on intentionally killing innocents. In drawing moral distinctions among cases with the same consequences, it directly challenges all forms of act consequentialism.

Distinguishing Intention from Foresight

So the DDE does a lot of good work for the absolutist. But there is a difficulty with the DDE, and it must be solved before absolutists can rely on it with confidence. The difficulty is that *we lack a clear basis for distinguishing between intention and foresight*. Without clarity on this point, the DDE will either fail to provide guidance about the morality of actions or will give us results that seem deeply mistaken.

Consider this challenge. Those who secretly abduct and carve up innocent people to distribute their organs could say that they intend only to save many innocent lives, and would be delighted if their innocent victims were (miraculously) to remain alive after the operation. Therefore,

they *don't* intend to kill their victims. They merely foresee their death. Thus the DDE does not condemn their actions.

It is hard to imagine someone saying this with a straight face. But explaining precisely what is wrong requires us to offer a sharp definition of intention. Further, this definition must clearly distinguish intention from foresight, and also reveal how harms that are intended can be wrong, while the very same harms, if foreseen, are morally acceptable. This isn't easy. Consider some prominent attempts:

- (A) You intend to do X = You want X to occur as a result of your action.

But the surgeon who is carving up the kidnapped victims may not want them to die. He may only want to save the lives of the many patients who need these organs to avoid a premature death. So with A, the surgeon does not intend to kill the abducted patients. But he surely does, and that makes A implausible.

- (B) You intend to do X = X is part of your plan of action.

Consider another variant on the trolley case. Here, the runaway trolley is heading toward five people, but there is no spur. The only way to stop the train is by pushing a huge bystander onto the tracks at the last minute. His bulk will stop the train—though he will surely die as a result. Here again, we save five at the cost of one. But this case leaves an awfully bad taste in the mouth.

Yet if I were to push this guy, I could deny that his death was part of my plan. My plan, as I formulated it to myself, didn't involve his death. My plan was limited, let's say, to pushing this man and to stopping the train. I'd be pleased if the man were to escape with only bruises. So, according to B, I didn't intend to kill the man. And that seems mistaken.

- (C) You intend to do X = You would regret it if X didn't occur as a result of your action.

Consider the last trolley case again. It seems clear that I intentionally killed the man I pushed to the tracks. But I would *not* regret it if he survived. Therefore, by C, I did not intend his death. Again, something has gone wrong.

- (D) You intend to do X = X results from your actions in a non-accidental way.

The problem here is that all merely foreseen results will now become results that we intend to produce. When the ER doctor gives the serum to her five patients, one patient will die. The doctor knows this. And it isn't an accident. Therefore, by this definition, the doctor intends that she die. And that, too, seems mistaken.

(E) You intend to do X = You must cause X if you are to achieve your goals.

In the second trolley case, it is false that the huge bystander's death must occur if I am to achieve my goals. All that *must* occur is that he stop the train with his body. And so, with E, I do not intend his death. And so the DDE does not condemn my action.

These aren't the only possibilities for defining what it is to intend to do something, but it's been a very difficult task for absolutists to provide a definition that manages to track the moral distinctions that we feel so confident about. So if you are a fan of absolutism, and the DDE, here is your task: clarify the distinction between intended and merely foreseen results, and do so in a way that shows why some intentional harms, just by virtue of being intended, are morally worse than harms that are foreseen. Perhaps it can be done. But it won't be easy.

Moral Conflict and Contradiction

For those who find ethical pluralism attractive, an invitation: Put together your preferred list of moral rules. Once you've completed the task, consider this question. What happens if these rules ever conflict with one another? If each of them is absolute, then moral conflict leads to contradiction. That is a very bad thing.

Consider a very simple case. Suppose that the correct list of absolute moral rules includes these two: Keep your promises. Don't deliberately harm innocent people. Now imagine that someone has done you a great kindness. As a gesture of thanks, you promise to help him whenever he needs it. He smiles, and then asks you to make good on your word by breaking into his rival's home and beating him senseless. The rival, let's suppose, is an innocent. He's done nothing to deserve such a beating.

In this case, you are morally required to keep your promise. But in doing that, you'd also be acting immorally, since keeping the promise means violating the rule against harming innocents. If you instead avoid hurting the innocent rival, you will be doing what is required of you. But

this would also be forbidden, since it would involve breaking your promise. Either way you go, your action is both morally required and forbidden. That is a contradiction.

It's a simple case, as I've said. But it is meant to illustrate a crucial worry for pluralist theories that endorse absolute moral rules. When such rules conflict, the theory yields contradictory advice. And that will sink the view, since contradictions are fatal flaws for any theory. The *Argument from Contradiction* summarizes this challenge:

1. If there is more than one absolute moral rule, then they are bound to conflict at some point.
2. If absolute rules ever conflict, then this generates contradiction.
3. If a theory generates contradiction, then it is false.
4. Therefore, any theory that endorses the existence of more than one absolute moral rule is false.

Premises 2 and 3 are true. They won't attract any challenges from philosophers. The real debate focuses on premise 1. If you believe that there is more than one absolute moral rule, then you have to show that the rules will never, ever conflict.

There is a way to do this. We will have to assume that absolute moral rules have a certain limitation: They forbid us from acting in specific ways, but they never require us to act in any way at all. We may never do certain things—rape, torture, and terrorize, for instance. These rules, and others, can be honored entirely through *inaction*. So long as all absolute rules can be obeyed by doing nothing, then absolute rules will never conflict.

This means that most of the familiar moral rules cannot be absolute, since they require us to act in certain ways—perform easy rescues, prevent disaster, protect our children from harm, keep our promises. If moral rules require us to do things, rather than merely refrain from action, then they might require us to do incompatible things. That is the kind of conflict that leads to contradiction. If, instead, all absolute moral rules can be satisfied merely by sitting on our hands and doing nothing, then contradiction can be avoided.

Is Moral Absolutism Irrational?

Many people think that intentionally killing innocent people is never morally acceptable. What could explain this? Presumably, the incredibly

important value of innocent human life. But if this is of paramount importance, then why shouldn't we kill an innocent person if we could save more innocent lives by doing so?

Optimists will say that such a thing could never happen. But it could. Suppose that you are a doctor in a remote community, and have just gotten word of a deadly plague in the area. There is no cure. You know the symptoms of the disease, and spot them in a newcomer to your village. You know that if any villagers come into contact with the infected man, those villagers will die—as will anyone who comes into contact with *them*.

You see the newcomer approach the village square, and can easily tell what is about to happen. You call out to the newcomer to stop, but either he doesn't hear or doesn't understand your warning.

Suppose that you are a crack shot. You now have a choice—gun this man down or allow him to go on his way, knowing that this will mean the death of a great many innocent villagers. This newcomer, of course, is himself an innocent. But if innocent life is so important, then why respect his life when it comes at the cost of so many more?

It seems irrational to defend an absolute ban on killing (or torturing or terrorizing), since we can sometimes *better* protect the underlying values of the ban by violating its terms. That shows that the ban should not be absolute. The charge can be put in the form of an *Argument from Irrationality*:

1. If perfect obedience to a rule can frustrate the purpose of the rule, then the rule is irrational.
2. Perfect obedience to any absolute moral rule can sometimes frustrate its underlying purpose.
3. Therefore, absolute moral rules are irrational.

This may look a bit like the Argument from Disaster Prevention, but there is an important difference. The charge here isn't that following absolute rules may produce a catastrophe. Rather, the complaint is that there is something fundamentally *inconsistent* about the absolutist position. It is tempting to defend absolute rules by claiming that they protect all-important values, such as the value of innocent life. But if these values can be better served by violating those rules, then the rules *should* be broken. And that means that they aren't really absolute, after all.

Absolutists have replied to this challenge by rejecting the argument's second premise. To see how they do this, consider again an absolute ban

on killing innocents. It's natural to think that the point of such a rule is to protect innocent life. But absolutists *deny* this. For if they granted it, then, as we have seen, their ban would be pointless, or worse, in any case where we could save more innocent lives by taking one of them. Certain absolutists think that we should never, ever kill another innocent human being—even if doing so will thereby spare many more innocent people from being killed.

But if the point of enacting a ban on killing innocents is not to save innocent life, then what is it? The answer is simplicity itself. The point of such a ban is to forbid you from killing innocent people. The only way to honor that demand is by not killing innocents. Of course, it may be (as in the example just given) that by respecting this requirement, you thereby open the door to a situation in which many more innocent people are killed. But the absolute requirement is not that you prevent the killing of innocents. It is that you kill none of them yourself.

In other words, the fundamental rationale for absolute moral rules is to forbid people from acting in certain ways, rather than to minimize the violation of these rules.

How can such a view be defended? The easy way is closed to us. The easy way tells us that innocent life is of the greatest importance, and so must be protected via an absolute rule. As we have seen, that way won't work, because this value can sometimes be better served by violating the rule against killing. That's what happens when we must kill one innocent person in order to save many others.

What we need instead is a defense of this key idea:

(A) We are always forbidden to act in certain ways, though we are not always required to prevent such acts from occurring.

Absolutists have found a way to defend claim A. It is called the **Doctrine of Doing and Allowing**.

The Doctrine of Doing and Allowing

At the finale of *Batman Begins*, Christopher Nolan's first movie about the comic book hero, Batman and his arch nemesis Ra's Al Ghul are having their climactic battle aboard an elevated train. The train is about to plunge hundreds of feet to its destruction. Batman has finally pinned his enemy and could deal the killing blow. But he doesn't. Instead, the last words that Ra's al Ghul ever hears are these: "I'm not going to kill you. But I don't have to save you."

Batman's words reveal a commitment to a view that absolutists have long relied on. This is the Doctrine of Doing and Allowing (DDA):

It is always morally worse to do harm than to allow that same harm to occur.

It is worse, for instance, to kill than to allow victims to be killed; worse to be a terrorist oneself than to allow others to perpetrate terrorism.

If it is always worse to do bad things than to allow them to occur, then we have a defense of the idea that absolute moral requirements apply only to what we do, and not to what we allow or fail to prevent. If the Doctrine of Doing and Allowing is true, we may be absolutely forbidden from killing, torturing, or raping, even though we are not absolutely required to prevent such behavior.

The DDA can therefore explain why there is something especially problematic about doing evil yourself. It can support the idea that you must keep your hands clean, even if doing so lets others cause even greater harm. The DDA directly supports principle A, just given, and is therefore essential to a defense of absolutism.

The Doctrine of Doing and Allowing thus allows us to deal with an especially tricky kind of moral dilemma—cases in which you are faced with a choice of doing something truly awful, or refusing to do so, knowing that your refusal will allow a less scrupulous person to take your place and do far greater harm.³

Consider the case of certain Nazi military officers who came to view their cause as deeply unjust. They had much to fear, for if their views became known, they would be shot. If they resigned or sought to run away and were captured, they would also be shot. Apart from such fears, however, another consideration kept some of them from leaving their posts.

These officers often correctly thought that if they didn't follow orders, then some true believer, with no inhibitions about carrying out the Nazi program, would replace them. And that would be even more disastrous.

By remaining in their positions, these officers were colluding with evil. They sometimes directly perpetrated it themselves. But they held the view, shared by many in situations like theirs, that the best way to try to destroy a morally bankrupt institution was to work from within. The problem, of course, is that one must often do evil to remain an insider. Immoral groups usually test the loyalty of their members, and this requires them to do all sorts of terrible things. Those who end up following orders often justify

their cooperation with evil by arguing that any alternative would bring about even greater harm. And they may well be right.

Together with moral absolutism, the DDA forbids certain actions, without requiring that you prevent others from doing them. On the safe assumption that Nazi atrocities were evil, the DDA would forbid German officers from committing them. It would require those officers to step down from their posts, even though, in doing so, they were signing their own death warrant, and effectively allowing others to do even greater damage.

We are sometimes forced to choose between doing actions that will stain our lives forever, and refusing to do so, knowing that our refusal will create an opportunity for other people to do their worst. In such cases, the DDA requires that we abstain from evil, even if this means that more people get hurt or that our own life is thereby placed at greater risk. For those who favor the DDA, some things are more important than minimizing harm to others, or even remaining alive. Preserving our moral integrity is one of them.

But is the DDA true? It may well be. Here are two cases that reveal its attractions.

Case 1: I am morally allowed to eat at a decent restaurant, spending thirty dollars on the dinner—even though, with that money, I could have prevented the death of three starving Somali children by sending a check to UNICEF. Most of us think it acceptable to spend the money on dinner, thereby allowing those children to die. Yet it would definitely be wrong to kill them directly.

Case 2: I am a soldier on the battlefield and see a fellow infantryman shot at my side. I can't tell whether his wound is fatal. He is groaning and asking for help. It would definitely be wrong to shoot that soldier. It would not be wrong to let him die, if I had to ignore his pleas in order to continue our attack or obey a command to retreat from the battlefield.

We are forbidden from doing something in both of these cases, though permitted to allow it to happen. The Doctrine of Doing and Allowing perfectly explains why this is so.

But some have had their worries about the DDA. Consider these two cases:

Case 3: I am a nurse charged with giving a bed-ridden patient a life-saving pill every four hours. I am an evil nurse, however, and when the appointed time comes, I stand by, pill in hand, and do nothing, thereby allowing the

patient to die. Suppose instead that I had obtained a poison pill and had given him that, rather than his real medication. He dies as a result. If the DDA is true, then giving the poison pill should be morally worse than failing to give the correct pill. Is it?

Case 4: I am a switchman who notices a runaway trolley speeding down the tracks, headed for one innocent person trapped in a narrow pass. I pull a lever and switch the tracks so that the trolley heads off on a spur. I do this because five people are on that spur and I want to kill as many people as I can. They all die. Now suppose, instead, that the runaway trolley is headed for the five rather than the one, and I just stand there, gleefully passing up an easy chance to switch the trolley to the track with only one person on it. The DDA says that my action is morally worse than my omission, that pulling a lever is morally worse than not pulling it—even though both cases result in the same number of deaths, and in both I have the same motive, intentions, and knowledge of the likely outcomes.

These are all very schematic, under-described cases. Some of them are quite outrageous. But there is a point to relying on examples like these. We can appreciate it by thinking about the nature of scientific experiments. The ones that are well designed often place two groups in identical circumstances, and change just a single variable, to test for its importance. For instance, we offer a control group a placebo, and then test a new drug by seeing how members of a comparison group fare when taking it. This allows us to isolate the drug's powers (or lack of powers).

The tests we conduct in ethics are thought experiments. We think of different situations that are identical in all respects but one, and then ask whether that difference makes any *moral* difference. That is what is going on in the cases just described. Each case presents two scenarios that are perfectly similar in every way but one. In the first scenario, the person *does* something very harmful, but in the second, the person *allows* that same harm to occur. Keeping everything else the same enables us to filter out any distracting, irrelevant details that could skew our responses to these cases. We can then determine whether the difference between doing and allowing, in and of itself, is morally relevant.

The problem for the DDA is that it sometimes seems, as it might in cases 3 and 4 (and other easily constructed examples), that the mere difference between doing and allowing has no moral importance at all. That is a serious challenge to the DDA. But of course your views on these cases may differ from mine, and you might feel strongly that doing the harm is, in each case, worse than failing to prevent it.

Even if that is so, however, there is another potential problem for the DDA: the distinction between doing and allowing is actually very hard to draw. Consider the classic case of "pulling the plug" on a patient whose life is being maintained by a ventilator. It seems that this is a matter of doing something—removing a tube or flipping a machine's switch. The doctor or nurse is not just standing there, stock still. But many have argued that removing such life support is merely letting nature take its course, i.e., allowing some ongoing sequence of events to continue on its way.

Before we can apply the DDA, we need to have a plausible way to distinguish doings from allowings. I don't know how to do this in a general way. But there are many easy cases, ones in which an action counts very clearly as doing something, or an inaction counts as obviously allowing something to occur. Perhaps this is enough to give absolutists hope that we can someday formulate a criterion that will enable us to distinguish doings from allowings once and for all.

The point of investigating the merits of the DDA is to examine whether the absolutist position can be defended against the charge of irrationality. Absolutists can defend against this objection if they can show why we must never commit certain harms, even if doing so would reduce the number of those harms. The DDA is meant to provide this piece of the puzzle. It tells us that doing harm is morally worse than allowing that same harm to occur. Whether the DDA is plausible will depend, in the end, on whether absolutists can supply a test for distinguishing between doing and allowing, and whether the DDA is immune to counterexamples of the sort that cases 3 and 4 seek to provide.

Conclusion

When I reflect on my own moral views, I find it hard to resist the thought that certain kinds of actions just should never be done. At the top of my list is murderous terrorism—deliberately trying to kill civilians, with the aim of inspiring fear and thereby furthering some political goal. When I vividly consider such behavior, I can't help but feel deeply repelled and outraged. I'm inclined to think that terrorism is always immoral.

But I know enough about myself to recognize my fallibility. My sense of outrage is not a surefire test of morality. Perhaps terrorism, in rare cases, is morally acceptable. Ditto for other familiar cases of repugnant behavior—rape, for instance, or torture. Whether such practices are always wrong depends on whether absolutists can adequately defend against the

Argument from Disaster Prevention, the Argument from Contradiction, and the Argument from Irrationality.

The worry about contradiction is the most serious, for any theory that generates contradictions is certainly false. So absolutists must show that their favored moral rules will never conflict. The only way to ensure this is to restrict absolute rules to those that require only inaction on our part. But why is it so important to refrain from doing things—especially if doing things will reduce the number of times that a moral rule is broken? The DDA and the DDE are designed to answer that question. The fate of absolutism hangs on whether these doctrines can be successfully defended.

Notes

1. See the introduction, pp. 12–14, for a more detailed discussion of the relevant attractions.
2. See chapters 9 and 10 for an extensive discussion of consequentialist ethics.
3. For further discussion of this matter, from a consequentialist point of view, please see chapter 10, p. 135.

Ethical Pluralism: Prima Facie Duties and Ethical Particularism

Ross's Ethic of Prima Facie Duties

Every moral theory we have considered thus far is absolutist. Most of these views are monistic, defending the idea that there is just a single absolute moral rule. But as we saw in the last chapter, some absolutists reject monism. They think that there are a number of moral rules that may never be broken. It's now time to take a look at another option. These are the theories that reject both monism *and* absolutism.

Such theories are pluralistic; they endorse the existence of at least two fundamental moral rules. And each of these rules is non-absolute; in some cases, it is morally acceptable to break them. W. D. Ross (1877–1967) was the philosopher who first developed this version of pluralism. He had a special term for these non-absolute rules. He called them principles of **prima facie duty**, and we will stick with that label in what follows.

A prima facie duty is an excellent, non-absolute, permanent reason to do (or refrain from) something—to keep one's word, be grateful for kindnesses, avoid hurting others, etc. As Ross saw it, each prima facie duty is of fundamental importance. None of these duties can be derived from one another, or from any more basic principle. Crucially, each prima facie duty may sometimes be overridden by other such duties. Though there is always good reason, say, to keep a promise or prevent harm to others, morality sometimes requires that we break a promise or do harm. Likewise for each of the other prima facie duties.

Ross was convinced that absolutism in all of its forms is implausible. He thought that pluralistic versions were bound to yield contradiction,

especially given the many fundamental moral duties that he defended (see the list in the next paragraph). And he rejected all forms of monism because they failed to recognize the diversity of fundamental moral considerations. For instance, while Ross accepted the utilitarian emphasis on doing good and preventing harm to others, he also agreed with Kant that justice was morally important in its own right, even if it failed to bring pleasure to others or to prevent their misery.

Ross identified seven *prima facie* duties, each of which is meant to represent a distinct basis of our moral requirements:

1. *Fidelity*: keeping our promises, being faithful to our word.
2. *Reparations*: repairing harm that we have done.
3. *Gratitude*: appropriately acknowledging benefits that others have given us.
4. *Justice*: ensuring that virtue is rewarded and vice punished.
5. *Beneficence*: enhancing the intelligence, virtue, or pleasure of others.
6. *Self-improvement*: making oneself more intelligent or virtuous.
7. *Non-maleficence*: preventing harm to others.

Ross made no claim to have provided a complete list. He allowed that there might be other *prima facie* duties. But each of these seven duties, he thought, definitely did belong on the list.

The term *prima facie duty* can be misleading. That's because these things are not really duties, but rather permanent moral reasons that partly determine whether an action really is, in the end, morally required. To say, for instance, that there is a *prima facie* duty of beneficence is to say that

1. there is *always* a strong reason to benefit others;
2. this reason is fundamental, and cannot be derived from any more basic reason;
3. this reason may sometimes be outweighed by competing reasons; and
4. if this reason is the only moral reason that applies in a given situation, then benefiting others becomes our all-things-considered duty—i.e., what we are really, finally morally required to do in that situation.

Focus for a moment on the first clause. It provides us with a way to test Ross's specific roster of *prima facie* duties. Suppose that there are situations

in which there is *no reason at all* to benefit others. If that were so, there would be no prima facie duty of beneficence.

I will let you do the testing yourself, because I am most interested in the general theory of prima facie duties, rather than in any specific version of it. Even if Ross's group of seven rules includes too much, or too little, this would not undermine the ethic of prima facie duties. What it would show (and this would certainly be important) is that Ross's own list was off-base. But a better list might make the cut.

So let's instead consider the big picture, and reveal the attractions, and difficulties, of the general model of morality that Ross advanced. First, as usual, the attractions.

The Advantages of Ross's View

Pluralism

The greatest attraction of the ethic of prima facie duties is its ability to accommodate our sense that there is, indeed, more than just a single fundamental moral consideration. To Ross, and to most of the rest of us, it does seem that the very fact of our having promised to do something generates *some* reason to follow through, even if keeping our promise fails to bring happiness, reward virtue, prevent misery, or anything else. That we have given our word is reason enough to do what we have promised.

But no one believes that promising is the only thing like this. There does seem to be something morally worrying, for instance, when someone repays a kindness with ingratitude—even if, in unusual circumstances, being ungrateful is the right way to go. Whether or not you agree with the whole of Ross's list, you may well sign on to the idea that fidelity and gratitude, at the very least, both possess independent moral importance. If you do, that is enough to force a shift away from monism.

We Are Sometimes Permitted to Break the Moral Rules

Ross's position also easily explains the widespread belief that the moral rules may sometimes acceptably be broken. There is always something to be said in favor of keeping a promise—but I should break my promise to meet a student for coffee if my daughter has a medical emergency and needs to be taken to the hospital. We all accept that there are circumstances in which it is morally acceptable to break a promise, allow harm to others, etc. Ross's theory straightforwardly explains this.

Moral Conflict

The ethic of prima facie duties also appears to make good sense of our experience of moral conflict. Duties conflict when they can't all be fulfilled. On absolutist views, such conflict yields contradiction. But Ross's theory easily avoids this.

Consider the case of a poor single mother whose child is too sick to go to school. The mother has a duty to report to work. By taking the job, she has in effect promised to reliably show up as scheduled. But suppose that she has just moved to town, has no friends or family there, and isn't allowed to bring her child to work. She also has a duty to care for her child, especially if no one else is available to do so. What should she do?

The Rossian can say of such a case that there is a conflict of prima facie duties. There is a strong case for showing up to work. There is a strong reason to care for one's child. Sometimes we can't do both. But no contradiction occurs, because we can distinguish between a strong standing reason (a prima facie duty) to do something and an all-things-considered, final duty to do it. When these final duties conflict—when we say, *in the end*, that you are absolutely required to show up at work and are also absolutely required to care for your child—then there is contradiction. Ross's view avoids this problem entirely.

Consider a district attorney who has evidence of a child molester's guilt. The problem is that the police obtained the evidence without a search warrant. Let's assume that the evidence is watertight. Not only that—there is, in this case, every reason to think that this man is going to repeat his crimes if he is released. The DA can doctor the paperwork so as to cover up the procedural error. He has a duty not to do so. But he also has a duty to prevent crime, and by tampering with the evidence a bit, he can do just that.

Perhaps this strikes you as an easy case rather than a hard one. If that is so, Ross has an explanation for that, too: though there is a prima facie duty, say, to obey the law and a prima facie duty to prevent harm, it is clear (if this really is an easy case) that one of these is stronger than the other. Recall that a prima facie "duty" is not really a duty, but rather a permanent, very strong reason to do something. Such reasons need not win the day in every case.

I'm not intent on defending a specific verdict in this example. If Ross is correct, the key thing is that context will determine just how important a prima facie duty is. The consideration at the heart of such a duty (promise-keeping, preventing harm, righting one's wrongs, etc.) is always morally

important. But it is not always morally *decisive*. That is precisely what distinguishes a prima facie duty from an absolute one.

Moral Regret

Another way in which Ross's theory very nicely handles moral conflict is in its view of moral regret. When moral claims conflict, and we can't honor them all, we think that it is right to feel regret at having to give up something important. Regret is evidence that something of value has been sacrificed. When prima facie duties conflict, and one takes priority over the other, the lesser duty doesn't just disappear. It is still has some weight, even though in the circumstances it is not as morally powerful as the conflicting duty. Regret is our way of acknowledging this forsaken duty, our way of recognizing that something of value was lost in the conflict.

Indeed, this provides us with a reasonable test for knowing what our prima facie duties are. The test is simple: there is a prima facie duty to act in a certain way only if it would always be appropriate to regret our failure to act that way. If there were nothing valuable about gratitude, for instance, then missing a chance to express it would not be a cause for regret. But it is. And that shows that there is *something* important about gratitude, even if it isn't *all-important*. That's just what Ross believed.

Addressing the Anti-Absolutist Arguments

Ross's view also provides a direct reply to two of the arguments that challenge the absolutist in the previous chapter. Consider the Argument from Disaster Prevention. This argument claims that any moral rule may be broken if that is what it takes to prevent a catastrophe. And therefore no moral rules are absolute. Ross naturally agrees with this. None of his moral rules is absolute. When the stakes are high enough, each of them may acceptably be sacrificed.

The Argument from Irrationality charges that absolutism is inconsistent, since the values at the heart of its rules can sometimes be better served by violating those rules. Ross can agree with this criticism as well. If we must break a promise in order to ensure that many more are kept, Ross can allow that this promise ought to be broken. The charge of irrationality stems from the absolutist claim that certain rules must be obeyed no matter the consequences. Ross rejects this claim.

But Ross also denies the push to consequentialism that lies at the heart of these two arguments. Both of them try to show that the results of our actions always determine their moral status. If obeying an allegedly

absolute rule yields disaster, or leads to greater violations of that rule, then we must reject the rule. We must achieve the best results, and since that sometimes means breaking the rules, this shows that the rules are not absolute.

Ross believed, of course, that the moral rules are not absolute. But he did not get there by assuming, with consequentialists, that our moral duty is always to maximize the amount of good in the world. Though he agreed with utilitarians that results are morally important, he denied that they are all-important. Doing justice, for instance, or improving oneself, is sometimes more important than doing what is optimific.

Indeed, to make his case against consequentialism, Ross had us imagine a situation in which we are faced with a choice. We will benefit person A or B. We can benefit person A a great deal, by fulfilling our promise to him. Or we can benefit person B just a very, very slight bit more, though we have made no promise to him. If we benefit B, we break our promise to A. Further, B has no expectation that we will benefit him. Ross thought it obvious that we ought to keep our promise to A, even though we would do just a bit more good by benefiting B.

This thought experiment convinced Ross of two things. First, that there is a *prima facie* duty of fidelity. There is always something morally important about keeping our word. And second, that consequentialism is mistaken, since this is a case in which one option (benefiting B) produces the most good, but is morally wrong. So, while Ross agrees with the conclusion of the two criticisms of absolutism here, he denies that we should be led to consequentialism as a result.

A Problem for Ross's View

In Ross's view, preventing harm is always morally important. Sometimes it is the most important thing you can do. But not always. Seeing that the guilty get their just deserts is also, and always, very important. If Kant is right, it always takes priority over preventing harm. If utilitarians are right, it never takes priority. If Ross is right, it sometimes does, and sometimes doesn't.

This leads us naturally to what may be the strongest objection to Ross's view. Ross denies that there are any absolute moral rules. So each moral rule may sometimes be broken. *But when?* The easiest way to answer that question would be to create a permanent ranking of the rules, by placing them in order from least to most morally important. Whenever a

lower-ranked rule conflicts with a higher-ranked one, the higher rule wins out, and determines our moral duty.

Ross rejected this strategy. He thought that there is no fixed ranking of the various *prima facie* rules, no permanent ordering in terms of importance. And he is not alone in thinking this. Though a ranking system is possible in principle, in practice no one has ever made it work. Sometimes it is morally more important to be grateful than to prevent harm. But not always. Sometimes it is more important to be honest with people than to spare them the hurt feelings that honesty may cause. And sometimes not. You get the picture.

And the problem: if we can't provide a fixed ranking of moral principles, then it isn't clear how we are to decide what to do when they conflict. That is because none of the *prima facie* duties has any kind of built-in moral weight. They are always important. But just how important? That depends on the specifics of the situation. Yet there are no guidelines that we can use from case to case to help us to know when a *prima facie* duty takes precedence over a competing duty. If a duty is sometimes, but not always, more important than another, then how do we know which one to obey when we cannot obey them both?

This is an extremely hard question. But before we can answer it, it seems we must answer an even deeper challenge: How can we know which *prima facie* duties are real and which are mere pretenders? For instance, is there a *prima facie* duty of honesty, or truth-telling? Many believe there is, but Ross omitted this from his list. How do we settle the matter, when people disagree about what which *prima facie* duties are genuine? It seems we first must answer that question and know the true list of *prima facie* duties, before we can get to the more specific question of how to strike an appropriate balance when *prima facie* duties conflict.

Knowing the Fundamental Moral Rules

Here is one of the hardest problems in ethics. How can we know what the fundamental moral rules are? The standard way of justifying a rule is not open to us here. We cannot cite a more general rule to back up the one in question. If the rule is really fundamental, then there are no deeper rules from which it derives its force. When we call such a rule into question, how can its correctness be defended?

Traditionally, there have been three strategies for dealing with this very hard problem: skepticism, **coherentism**, and appeals to **self-evidence**. Let's consider them in turn.

Skepticism

The skeptical strategy denies that we can know what the fundamental moral rules are. We can't have such knowledge because we can't know that *any* moral rule is correct. That's a pretty extreme claim. How can such skepticism be defended?

Fans of this strategy say that every moral rule needs some defense—no rule can stand alone, no moral rule is beyond question. True, we can defend a rule by introducing another rule to support it. But this supporting rule itself can be questioned. We can defend *it* by bringing in a third rule. But that third rule can also be questioned. The worry is that this sort of questioning can go on *forever*. And if that is true, then we can't know that any moral rules are correct.

Consider: What's wrong with teasing children? It humiliates them. What's wrong with humiliating children? It is emotionally painful. What is wrong with imposing emotional pain? It is a form of harming people. What's wrong with harming people? And so on. In the real world, such questioning always comes to an end. But that's not because we have arrived at a secure stopping point. It's just because people get impatient, or exhausted, or stumped. But without those shortcomings, the questioning would never stop. Any moral claim you put forward as a defense of another moral claim can itself be questioned, and on, and on... forever.

Philosophers call this unhappy kind of situation an **infinite regress**. It is a never-ending chain of questions and answers. The problem is that since there is no stopping point, none of the claims along the way can be justified. We are justified in believing in a moral rule only if we can defend it. A defense requires that we be able to have an anchor, some moral rule that can end the regress. But since every single moral rule requires support, there is no such end. If the line of questioning can go on forever, then we are not justified in believing in any moral rule, no matter how clear or obviously true it seems to us. And that means that we can never know that any moral rule is correct.

Coherentism

A second option is more optimistic. In this view, known to philosophers as *coherentism*, we can justify any moral claim (including a fundamental moral rule) by showing that it receives support from, and lends support to, a large number of our other beliefs. If a belief coheres well with many other things we already believe, then it is to that extent justified. And if that belief is true, then (barring some accident) it will also count as knowledge.

So, for instance, we could regard as well justified the belief that it is *prima facie* wrong to harm someone. That's because this belief is going to support, and be supported by, a lot of our other moral beliefs. If this moral rule is correct, and coherentism is true, then we can know that harming others is *prima facie* wrong.

There is a very old worry about coherentism—namely, that it approves of **circular reasoning**. Circular reasoning amounts to defending some belief by a set of other beliefs whose justification ultimately traces back to the original claim in question. Really, it amounts to using a belief to justify itself, and that is very suspicious. According to coherentism, a belief will be justified only if it is supported by many other beliefs. But in order to support the original belief, these other beliefs have to be in good standing. They also need to be justified. And how will that happen? By their receiving support from other beliefs, including, ultimately, *the original belief in question*. And that is circular reasoning, which (according to critics) never justifies anything.

After all, no one thinks that the following line of reasoning justifies anything:

- Abortion is immoral.
—Why?
- Because it kills a fetus.
—Why is killing a fetus immoral?
- Because that would be aborting it, and abortion is immoral.

Abortion may be immoral, but this reasoning can't show it so, and that is because the original claim in question is used to defend itself. That is circular reasoning.

Coherentists reply that if the circle is large enough, the beliefs included in it may indeed be justified. When a belief partly supports many other beliefs, which support many others, which support many others, some of which are supported by the initial belief in question—what we have here is a broad network of mutually support with a wide variety of beliefs reinforcing others. For some especially well-established beliefs (e.g., that there is a physical world; that I have been alive for more than one day; that there is a *prima facie* duty not to cause intense physical suffering), there may literally be hundreds of other beliefs that support them, and are supported by them. This kind of mutual reinforcement creates a powerful web of beliefs; each strand (i.e., each belief) is justified precisely because of its reinforcing links with many other strands.

There is a great deal of debate within philosophy about the merits of the coherentist project. Rather than pursue the matter, let me note simply that coherentism, if true, makes Ross's theory quite compelling. A great many of our moral beliefs depend on the idea that there are always excellent reasons to keep our promises, avoid harming others, do justice, etc. And these prima facie duties also receive a lot of support from other beliefs, including many of our beliefs about what is morally important in specific situations. There is a broad network of mutually supporting beliefs that contains Ross's principles of prima facie duty.

Further, the thought that these duties have different weight in different contexts—another essential feature of Ross's view—also receives strong support from our considered moral beliefs. So if coherentism is correct, we may well have knowledge of many fundamental moral rules.

Self-Evidence

But this wasn't the picture of moral knowledge that Ross preferred. He regarded his seven principles of prima facie duties as *self-evident*. A claim is self-evident just in case it is true, and adequately understanding it is enough to make you justified in believing it. Self-evident truths are those that you are justified in believing on the basis of careful reflection alone. If you think hard about such claims, and come to believe them as a result, then you will have knowledge.

I think that there are some self-evident claims. Here are a few:

- All bachelors are unmarried.
- If Alice is taller than Bob, and Bob is taller than Charlie, then Alice is taller than Charlie.
- Anything that happened a decade ago occurred prior to today's events.
- Uncles have (or had) siblings.
- The sum of any two odd numbers is even.

Some of these claims are just obvious; others may take a bit of time to sort out. Self-evident claims need not be obvious. What is crucial is that careful reflection is all it takes to know them.

Suppose that some moral rules are self-evident. Then we have a way of calling a halt to an otherwise infinite chain of moral questioning. Our stopping points will be self-evident claims that require no further justification.

If Ross is right, then each of his seven principles is self-evident. We can know them just by thinking about what they really stand for. If we can

rid ourselves of bias, hasty judgments, overemotional involvement, we will be convinced that there is always something right about keeping promises, preventing harm, doing justice, showing gratitude, etc.

Self-Evidence and the Testing of Moral Theories

Ross thinks that his theory of prima facie duties, and his confidence in their self-evidence, is in deep harmony with common sense. And as he sees it, this is a great benefit of his theory. We should not overturn the biddings of common sense just because it conflicts with a pet theory. Ross used the example of beauty to establish this point. For instance, we should not abandon our strong belief in the beauty of the *Mona Lisa* just because some philosophy of art declares that only Impressionist paintings or medieval altarpieces are really beautiful. We should give up the theory before giving up on our deepest, most secure beliefs.

What is true of our artistic judgments is also true of our moral ones. We can see how this plays out by considering Ross's rejection of consequentialism. Ross was quite clear-eyed about how tempting consequentialism can be. But he insisted that it was fatally flawed because it failed to appreciate the variety of fundamental moral concerns. Consequentialism imposes order, system, and a unifying principle onto our moral thinking. But we must resist such charms, because they conflict with our deepest beliefs about what is truly morally important. Our confidence in the independent value of promise-keeping—or justice, or repairing our wrongs—should not be held hostage to a theory's demands.

If Ross is right, we use our deepest commonsense beliefs, some of which will be self-evident, as the way to test moral theories. Our self-evident beliefs have a kind of priority in moral thinking. It isn't as if each moral belief we have is beyond scrutiny. Far from it. Some of our moral views, perhaps even our most cherished ones, may have to go, once we see that they conflict with beliefs that are even better justified. Still, the data of ethical thought, as Ross puts it, are those moral beliefs that have survived very careful reflection. Self-evident principles are where our moral thinking must begin. They are what moral theories must account for. These basic beliefs are to be given up only if we can show that they can't all be true.

To the extent that a moral theory cannot make room for such beliefs, it is the theory that must go. This was Ross's diagnosis of both consequentialism and Kantianism, for instance. They both understood morality too narrowly, as limited to a single fundamental moral rule. He thought that

careful reflection would show us that there are at least seven such rules—none of them absolute.

Ross acknowledged that his view offered little comfort to those who did not agree with his seven principles. But he was unapologetic. To someone who thought about justice, for instance, and failed to see its moral importance, Ross could do only one thing. He would invite that person to think more carefully about what justice really is. This can be done in many ways. We can offer the person examples to consider; draw analogies to cases that reveal the importance of justice; distinguish justice from other, possibly related, notions; ensure that particular beliefs opposing the importance of justice are not based on error. But suppose the person remains unconvinced even after all of this further reflection. According to Ross, moral discussion now comes to an end, and the only verdict to render is that this person is mistaken. Nothing you can say will show him that he is wrong.

That may strike you as closed-minded, but two things can be said in Ross's defense. First, what are the alternatives? Why must it always be possible to offer something more in defense of one's favored claims? If justification ever *does* stop somewhere, in ethics as in other areas, then once we have reached that stopping point, all that could possibly be done is to invite the doubters to reconsider. If, instead, justification is never-ending, then we are back to an infinite regress, which justifies nothing at all.

Second, we should consider the possibility, in *non-moral* contexts, of finding ourselves without any support for a claim that we rightly continue to believe. For instance, there may be nothing you can say that will convince a member of the Flat Earth Society of his mistake; no way to convince someone who believes in vampires that he is wrong; no clear path to show a stubborn person that nothing really can be all pink and all blue at the same time. You may be justified in your beliefs even if you can't always convince those who disagree with you. That holds for moral as well as for non-moral beliefs.

Knowing the Right Thing to Do

Even if our prima facie duties are self-evident, we are still faced with the problem of knowing what to do when they conflict. And Ross has very little to say here, except that we can never be certain that the balance we strike is the correct one. Ross acknowledged that our actual, all-things-considered moral duty on any given occasion is *not* something that is self-evident. We may feel very strongly about certain cases; indeed, most moral situations

are easy and straightforward ones, which we never give a second thought to. Still, there is no definite method for guiding us from an understanding of the *prima facie* duties to a correct moral verdict in any given case.

We must start our moral thinking about specific situations by understanding the kinds of things that can be morally important. This is a matter of clearly grasping the *prima facie* duties. These tell us what to look out for. Has a promise been made? A wrong been done? Is there an opportunity for self-improvement here? And so on. But once you answer such questions, you're on your own. You must bring your experience and insight to bear on the details of a given case. The bad news is that there is no fixed or mechanical procedure that tells us how to do this.

This can be very dissatisfying. There are several aims of moral theory, and one of them, surely, is to offer advice on deciding how to live. Ross denies that there is any general rule to follow in order to provide answers here. What a letdown.

But again, there are a few things we might say in order to make this a bit easier to swallow. First, perhaps the idea of a comprehensive moral decision procedure, one that can be consulted to provide definite answers to all moral questions, is not so plausible. When faced with puzzling ethical questions, we may *want* a concrete set of guidelines to help us along. But do we really believe that there is such a thing? Familiar options are fraught with problems. (Consider the difficulties of the golden rule, the *What if everyone did that?* test, or the principle of universalizability, for instance.) Perhaps the best explanation of this is that we are looking for something that does not exist.

Second, Ross's theory is not the only ethical view that faces this problem. Indeed, even consequentialism encounters this difficulty. If a consequentialist identifies more than one thing as intrinsically valuable, then it will be unclear what to do when we can maximize one value without maximizing the other. For instance, if happiness and fairness are valuable in their own right, then what should we do when faced with a choice between maximizing happiness (at the expense of fairness) or maximizing fairness (at some cost to happiness)? Things get even trickier if there are three or more intrinsic values. Yet even if there is only one such value, consequentialism still faces a problem. That's because it has always issued *two* fundamental moral rules: *Maximize goodness*, and *Minimize badness*. But what if you can't do both? Consequentialists don't assign permanent priority to one over the other. And it isn't clear how they would do so, in any event.

Finally, the absence of a decision procedure for arriving at conclusions is actually the *default* situation across all areas of thinking (except mathematics and its associated disciplines). For instance, scientists faced with a conflict between their data and some favored theory have no uniform method for determining whether to modify the theory or rethink their data. Further, even when the data are uncontroversial, selecting the best theory to account for it is anything but a rote, mechanical undertaking. Scientists must rely on good sense, too, since choosing which theory to believe is a matter of balancing the virtues of the competing theories. There is no precise rule to tell a scientist how to do this.

There are many theoretical virtues: parsimony (employing fewer assumptions than competing theories); conservatism (preserving as much as possible of what we already believe); generality (explaining the broadest range of things); testability (being open to experimental challenge and confirmation); and others. Suppose that one theory is more parsimonious and also more conservative, but another theory is more general and more testable. Or suppose that one theory is far more conservative than any competitor, but is also somewhat less general, and a fair bit less parsimonious. Science does not offer us any kind of definite procedure for identifying the better theory. Sometimes it is just obvious that one theory is better or worse than another. But in close cases, scientists have no alternative but to use their judgment.

And that is precisely our situation when it comes to morality. There are many easy cases, though these rarely get our attention, since they don't call for any hard thinking. It's the difficult cases—where different options each respect some *prima facie* duties, but violate others—that require judgment. We can never be sure that we've exercised good judgment. We may be unable to convince ourselves, much less our opponents, that we have landed on the right answer to a hard ethical question. The lack of guidance we get from Ross's view of ethics can leave us feeling insecure and unsettled. That is regrettable. But it may also be inescapable.

Ethical Particularism

The ethic of *prima facie* duties offers a serious challenge to both absolutism and to monism. Yet there is a view, known as **ethical particularism**, that is an even more extreme challenge. Particularists reject absolutism. They reject monism. They also deny the existence of any *prima facie* duties.

Recall that a central feature of such duties is that they represent moral reasons that are always important. Every time you do wrong, for instance, there is excellent reason to repair the damage. There is, without exception, something to be said for keeping the promises you've made. True *prima facie* duties point to features that are always morally relevant.

Particularists deny that anything meets this description. As they see it, something's moral importance depends entirely on context. If they are right, then sometimes there is nothing good *at all* about keeping a promise, or benefiting someone else, or preventing harm to others. The moral value of such things depends entirely on the details of the case. We have to consider all of the features in a given situation before we can know the moral contribution that any one of them makes.

Prima facie moral rules claim that certain features (e.g., promise-keeping, self-improvement, preventing harm) are always morally important. Absolute moral rules state that certain features are not only always morally important, but also morally decisive—these features settle the matter of our moral duty once and for all. Particularism, then, is the view that *there are no prima facie or absolute moral rules*.

That sounds pretty drastic. Certainly, particularism occupies one end of the spectrum of moral theories, with monistic, absolutist theories (such as ethical egoism or act utilitarianism) at the other. If particularists are right, then morality is entirely particular to specific situations; there are no moral rules at all to help us navigate our way in the world.

Consider an example of particularist thinking: It is often important to keep our promises, but not always, and if that is right, this shows that Ross was mistaken to think that promise-keeping is a *prima facie* duty. A promise made by a hostage to her kidnapper carries no weight. Nor is there anything good about returning a promised weapon to its now obviously homicidal owner. If particularists are right, there is sometimes no reason at all to keep a promise.

Particularists think that every *prima facie* rule is subject to this sort of criticism. For each of Ross's *prima facie* duties, and for any others, particularists will try to offer counterexamples to the rule. A *prima facie* rule says that X (self-improvement, doing justice, etc.) is always morally important; particularists will try to come up with cases in which X has no moral value at all. If they are right, nothing possesses any fixed moral importance. Whether something has moral value always depends on the other features in a situation.

Particularists often use non-moral examples to soften us up to their core idea. The intense coloring of a Turner painting is an essential part of what makes it so beautiful. But such colors would ruin a piece by Whistler. The paint drips that make Jackson Pollock's work so interesting would completely spoil a Monet. Lesson: whether something makes a difference to the beauty of an object depends entirely on context. There are no rules of beauty, because there are no features that always enhance (or spoil) the beauty of an object.

That a person told me something is sometimes a reason to believe it. And sometimes not. That an act is against the law is usually a reason to refrain from doing it. But sometimes it is a reason to perform it (say, when engaging in civil disobedience to protest injustice). That I want something is sometimes a reason for me to do it. But not always (consider an addict's desires). Here we have cases in which things sometimes but not always count as a reason. When this is so, the features in question have no place in either a *prima facie* or an absolute rule.

Particularists rely on such examples to show us that theirs is the ordinary view when it comes to non-moral matters. And yet when applied to morality, it has met with great resistance from philosophers. There are three primary sources of opposition.

Three Problems for Ethical Particularism

Its Lack of Unity

This criticism, and the next, should be familiar from our discussion of *prima facie* duties. Monists attacked Ross's theory because, in the words of one critic, it offered us nothing more than "an unconnected heap of duties." Monists felt uncomfortable with the idea of several fundamental moral rules, rather than just one. Indeed, as we have seen, philosophers are strongly tempted to look for a certain sort of moral theory—a body of unified, systematically interconnected claims that stem from a single fundamental truth. Ross's theory was a disappointment in this respect; particularism completely dashes such hopes. In the particularist picture, the moral realm is hugely complex, and there are no moral rules at all to help guide us on our way.

But this cannot be a decisive criticism of particularism. We might hope for simplicity and elegance, but the moral realm may be much messier than we had thought. We can't assume from the outset that monism is true, and then criticize particularism for failing to live up to its standards.

Whether the moral realm is neatly ordered, as absolutists believe; whether it is somewhat structured, as Ross argued; or whether it is highly disordered, as particularists insist, is a matter that can be settled only after a great deal of further moral debate.

Accounting for Moral Knowledge

The second criticism leveled against particularists is that their view provides us with no guidance for gaining moral knowledge. Ross's theory, as we know, came in for criticism on this front, since he failed to offer general advice about how to balance *prima facie* duties when they conflict. But particularism takes this worry to a new level, since Ross, at least, was able to instruct us on what to look out for in all of the cases we might face. The *prima facie* rules cannot by themselves settle what is actually right and wrong, but they serve as useful signposts to indicate the features that are relevant in discovering our moral duty.

By contrast, particularists tell us that anything that has once been a force for good may, at other times, be either morally neutral or positively bad. There is no way to know in advance how things are going to play out. Nor is there any method to follow that can clue you in to the correct moral verdict in a given situation. We don't have any rules to tell us what is morally important. And we lack rules to tell us how to figure out our moral duty in specific cases.

If the moral importance of everything depends on context, then there is no general roadmap to follow for those who want to know their moral duty. Indeed, particularists can offer almost no such advice at all, other than some very broad tips (take careful note of details, don't confuse self-interest with morality, get the facts straight, etc.). The particularist can borrow something close to Ross's story about how we gain knowledge of right and wrong. That's not much of a story, admittedly, and particularists are in one respect worse off than Ross, since they deny the existence of even *prima facie* moral rules. Still, though Ross's account of how to learn the right answer in moral cases is quite thin, there are several replies on his behalf that we have seen, and particularists can make use of those replies as well.

Basically, they must insist that our moral knowledge comes only through a comprehensive appreciation of all of the relevant features of a situation. Just as we can take in the beauty of a canvas only by noticing how each distinct feature plays off the others, so too we can detect the morality of an action only by taking careful note of all of its important features and

their interplay with one another. The only thing particularists add to this Rossian picture is this: which features are relevant in any given situation cannot be known in advance, since no features possess permanent moral importance.

In many ways, this picture is disappointing, since it fails to supply a general blueprint for gaining moral wisdom. Yet this is a serious failing only if there *is* such a blueprint. There may be. But that remains to be seen.

Some Things Possess Permanent Moral Importance

The deepest critique of particularism is that its central claim—that there are no features that are always morally important—is mistaken. If it is, then there are at least some *prima facie* moral rules. Whether there are any absolute rules depends, of course, on whether there are any features (e.g., maximizing happiness, being commanded by God) that are morally decisive in every possible context.

To undermine particularism, we would have to provide examples of features that are always morally important. Perhaps Ross managed to do this—even if promise-keeping, for instance, fails to make the cut, perhaps other of his *prima facie* duties fare better.

Consider doing justice. Relying on the “regret test” mentioned earlier (see p. 224), it seems that there is always some reason for regret when we commit injustice, even if injustice really is the way to go in a given case. But I don’t want to rest anything on my view of this matter. That’s because we can leave justice aside, and still show that particularism is in trouble. There seem to be a number of *prima facie* duties, even if they are not the ones that Ross himself favored.

For instance, I think that it is *prima facie* wrong to humiliate people, to hurt others strictly for the pleasure it gives you, to intentionally kill an innocent person who wants to live, to betray a friend’s trust, to knowingly violate an oath because of greed, and to commit rape. Maybe none of these represents an absolute rule; perhaps, in unusual circumstances, it can be morally acceptable to commit each of these acts. But there would always be something to regret in doing so, and that is good evidence that there is a *prima facie* duty not to do such things.

This list is certainly not complete. But if my list (or a better one that you can construct) is plausible, then particularism is mistaken. Morality would have more order and structure than particularists allow. They claim that we can never know in advance, independently of context, whether

something is morally important. If there are any *prima facie* moral rules, that claim is false.

Conclusion

The ethic of *prima facie* duties has a lot of things going for it. It is pluralistic, and so rejects the idea that the whole of morality can ultimately be explained by a single moral rule. It rejects absolutism, and so explains why it is sometimes permitted to break legitimate moral rules. It easily handles moral conflict without falling into contradictions. It offers an important role for regret in thinking about what is morally important. And it nicely handles some of the most difficult arguments designed to undermine moral absolutism.

Yet like all of the moral theories we have discussed, Ross's view is not without its problems. Perhaps the hardest of these concerns the question of how we can know what to do in particular situations. Since there is no permanent ranking of the *prima facie* rules, and no precise method for knowing how to strike a balance when the *prima facie* rules conflict, this leaves us with very little guidance for gaining insight into what morality actually requires of us.

Ross also has the worry of explaining how we might gain knowledge of his *prima facie* rules. I think that the question of how we can know the fundamental rules of morality is very, very hard. And Rossians can take some comfort in the knowledge that *every* moral thinker shares this problem. I sketched two general strategies (plus one skeptical response) that might be of some help here. But this of course offers only the very beginning of the needed discussion.

As we have seen, those who resist both absolutism and monism need not go in for *prima facie* duties. They may embrace ethical particularism, which denies that there are any moral rules at all. In this view, no features are always morally important, much less always morally decisive. Particularism is a bold thesis. But despite its boldness, it does seem that there are some features that are indeed permanently morally important. And if that is so, then there are some *prima facie* rules after all.

The particularist can still level a challenge to those who think that morality is rule-based. Imagine that after a great deal of thought, we were able to identify a dozen *prima facie* duties. It might still be the case that *most* of the morally important features we encounter in our lives were not mentioned in any of those rules, because these features are only sometimes, and not always, morally relevant.

For instance, suppose that my earlier criticism of promise-keeping was on target, and that there really are cases in which nothing at all favors our keeping a promise. Still, the fact that we promised to do something usually *is* morally important, even though (on the present assumption) there is no absolute or prima facie moral rule that tells us so. Likewise for telling the truth—there is ordinarily a strong reason to be truthful, even if there are some cases where there is no moral value at all in telling the truth.

If morality is really like this, where many features that are morally important are not always so, then we have a kind of hybrid view, a mix of Ross's theory and particularism. Some types of action possess permanent moral relevance, as Ross said; others, perhaps most, do not. If that were so, then we would be faced with a moral world that was *far* less simple and unified than the one presented by monists and absolutists. Is that our world? That is for you to decide.



Virtue Ethics

What sort of person should I be? An answer to that question provides some of the most vital information you can ever have. And yet none of the ethical theories we have examined thus far does much to address it. Each will say: You ought to be the sort of person who... maximizes happiness, or treats others with respect, or adheres to rules that free and equal people would endorse, or honors absolute rules. That's a pretty thin sort of answer.

To see what might be missing, consider what we might say of a police officer who obeys the law, but only begrudgingly. He always does the minimum required of him. If he could get away with it, he would extort money from business owners on his beat, brutalize prisoners, doctor evidence. He doesn't do any of these things, for fear of getting caught and punished. So he hasn't actually done anything wrong.

If we consider only what this man has or hasn't done, we will be missing a large part of the ethical picture. To fill things out, we must consider the kind of character he has. It's not a good one. We don't admire such a person; we don't want our children growing up like him. His manner is lazy, abusive, and untrustworthy, even if his conduct is satisfactory. If we keep the focus on whether he has done his duty, there is nothing to criticize him for. But that just shows that the focus needs to be shifted. We need to be less preoccupied with ideas of moral duty, and concentrate much more on ideals of character. Following that advice leads us directly to a consideration of **virtue ethics**.

All of the moral theories we have reviewed thus far share a common assumption. The assumption is that the moral philosopher's primary task

is to define the nature of our moral duty. In this view, *What should I do?* is the crucial moral question. Once we have an answer to that, I can know what sort of person I should be—namely, the sort who will do my duty as reliably as possible.

But what if we approached ethics from a different starting point? What if we began by considering what makes for a desirable human life, examining the conditions and the character traits needed to flourish? Rather than begin with a theory of moral duty, we would start with a picture of the good life and the good person, and define our duty by reference to these ideals. That is precisely what virtue ethics recommends.

Virtue ethics is not some single theory, but rather a family of theories that can trace its history (in the West) to the philosophy of the ancient Greeks. Aristotle's *Nicomachean Ethics* has had the greatest influence in this tradition, and remains a primary inspiration for most who work in it. Aristotle's book develops most of the major themes that even today define the virtue ethical approach to the moral life. Let's consider some of the most important of these themes.

The Standard of Right Action

Virtue ethics insists that we understand right action by reference to what a virtuous person would characteristically do. To put it a bit more formally,

(VE) An act is morally right just because it is one that a virtuous person, acting in character, would do in that situation.

According to virtue ethicists, actions aren't right because of their results, or because they follow from some hard-and-fast rule. Rather, they are right because they would be done by someone of true virtue. This person is a **moral exemplar**—someone who sets a fine example and serves as a role model for the rest of us. The ideal of the wholly virtuous person provides the goal that we ought to aim for, even if, in reality, each of us will fall short of it in one way or another.

Virtue ethics is actually a form of ethical pluralism. Though there is a single ultimate standard—do what the virtuous person would do—there are many cases where this advice is too general to be of use. At such times we need a set of more specific moral rules. Virtue ethics can provide these, too. For each virtue there is a rule that tells us to act accordingly; for each vice, a rule that tells us to avoid it. So we will have a large set of moral rules—do what is honest; act loyally; display courage; deal justly with others; show wisdom; be temperate; avoid gluttony; refrain from infidelity;

don't be timid, lazy, stingy, or careless; do not show hostility; free yourself of prejudice, etc.

When these rules conflict, how do we know what to do? We should follow the lead of the virtuous person. True, it won't always be easy to know this. There will inevitably be disagreement about who counts as being virtuous, and about the actions such a person would pursue. But this needn't cripple us. There is lots of room for critical discussion about who is virtuous and why. In the end, we may have to agree to disagree, since there may be no way to convince someone whose moral outlook is fundamentally opposed to our own. Someone who has been raised to idolize Hitler or Stalin is going to have a skewed moral vision, and there may be no way that we can convince him of his error. Virtue ethicists deny that this undermines their ultimate moral standards. It just shows that some people may always be blind to them.

Moral Complexity

Many moral philosophers have hoped to identify a simple rule, or a formal procedure, that could tell us precisely what our moral duty is in each situation.¹ What's more, this rule or procedure could be reliably used by anyone, so long as he meets a minimal degree of intelligence. A classic example of this is the golden rule. Even a five year-old can apply this test.

Virtue ethicists reject the idea that there is any simple formula for determining how to act. At the beginning of the *Nicomachean Ethics*, Aristotle cautions that we must not expect the same degree of precision in all areas of study, and implies that morality lacks rules and methods of thinking that are as precise as those, say, in mathematics. When it comes to morality, we must be content with general principles that allow for exceptions.

Virtue ethicists have followed Aristotle in this thought. To them, ethics is a complex, messy area of decision-making, one that requires emotional maturity and sound judgment. One of the *problems* of the golden rule, for instance, is that even a child can use it with authority. Aristotle thought it obvious that even the most perceptive children are still far short of true moral wisdom.

Virtue ethicists sometimes invite us to appreciate the complexity of morality by having us imagine a moral rule book. The book would contain all the true rules of ethics, and all of the precise methods for

applying them. It would state when exceptions were called for and when they were forbidden. It could be applied in a mechanical way, without any need of judgment. Is this a real possibility? Not likely, according to virtue ethicists. Morality is not like geometry or civil engineering. We have moral rules of thumb that can help us in most situations. But strict obedience to such rules is bound to lead us into error. And the rules, of course, will sometimes conflict. What we need in all cases is a kind of sensitivity. It is something very different from a rote application of pre-set rules.

This does not mean that everything is up for grabs in ethics. The precision of a discipline is one thing. Whether its principles, methods, and results are merely a matter of opinion, each one as good as the rest—that is quite another. Morality may be an imprecise discipline, but that does not mean that each person's moral views are as plausible as another's. Aristotle and most of his followers believe in objective standards of morality (those that are true independently of personal feelings or opinions). Whether they are right about this is something that we consider at great length in the final part of this book.

Moral Understanding

As virtue ethicists see things, moral understanding is not just a matter of knowing a bunch of moral facts. For if it were, then a child prodigy might be one of the morally wisest among us. As we have seen, virtue ethicists deny this possibility. Imagine turning to such a child for advice about dealing with difficult co-workers, or helping a drug-addicted friend through recovery, or determining the best way to break off a relationship.

Moral understanding is a species of practical wisdom. Think of some familiar kinds of practical wisdom—knowing how to fix a car engine, skillfully playing an instrument, being the sort of manager who knows how to inspire her workers and prevent petty office quarrels. Such knowledge does require an understanding of certain facts. But it is much more than that. We all know people with plenty of book smarts and very little in the way of good sense. Moral wisdom is a kind of know-how that requires a lot of training and experience. What it doesn't require is a superior IQ or a vast reading list.

We need experience, emotional maturity, and a great deal of reflection and training in order to acquire moral wisdom. We have to know how to read people, to be familiar with the sorts of troubles people can

fall into, to understand the kinds of personality issues that can prevent us from flourishing, to have a keen ability to pick up on social cues. There is an art to being a morally wise person. It takes intelligence, creativity, and sensitivity. We can't learn such things only from books.

One way to reinforce this idea is to understand the crucial roles that emotions play in moral understanding. There are three that are especially important.

1. *Emotions can help us to see what is morally relevant, by tipping us off to what matters in a given situation.* Fear can signal danger; compassion can tell us that someone needs our help. It's no use knowing that you ought to help those in need if you always walk around with blinders on, never aware of the struggles or potential discomfort of others. The person with the virtues of compassion, sympathy, and kindness will see things that others miss. Our emotions, when they are well-trained, reliably alert us to the morally important features of our lives.

2. *Emotions can also help to tell us what is right and wrong.* If we are virtuous, the anxiety we feel when considering certain actions is excellent evidence that these actions are immoral. We often feel that certain paths are simply off limits, or that other things definitely must be done, before we have a good intellectual account of why this is so. The pride of a good man is reason to think that he has done well. His anger is a reliable indicator that someone has done wrong.

3. *Emotions also help to motivate us to do the right thing.* They support and reinforce our thoughts about what we ought to do. Knowing the right course of action is one thing; following through is another. The morally wise person will have an easier time of things here, because her emotions will be in harmony with her understanding of what morality calls for. Unlike a weak-willed person or someone who manages to control her inappropriate impulses, the morally wise person wholeheartedly does what is right. She is relatively free of inner conflict and takes pleasure in doing the right thing.

Moral wisdom is an extremely complicated kind of skill. It does require knowledge of the way the world works, but it demands more than that. We must have a great deal of emotional intelligence as well. Virtues, which all require moral wisdom, therefore also require a combination of intellectual and emotional maturity and understanding. A person with only a crude appreciation for life's complexities, or a blank emotional life, is bound to be morally blind. Virtue ethics perfectly explains why that is so.

Moral Education

Virtue ethicists, again following Aristotle, believe that moral understanding can be gained only through training, experience, and practice. True, some people are by nature more generous or kinder than others. Yet an impulse in these directions is not enough. Without wisdom, these traits will only occasionally lead to appropriate action. We shouldn't always give to others or tend to their needs. They may be engaged in evil projects. They may need tough love, rather than indulgence. The wise person will know when to give, and when to withhold.

So virtue is not inborn. It takes time to acquire. And it also takes the right sort of environment and teachers. Indeed, Aristotle thought that whether we are virtuous or not is partly a matter of luck. If we grow up in a corrupt society with terrible role models, then we may lack the opportunity to develop virtues. We can't become virtuous overnight, and we can't do so without good teachers. The most important elements of moral education occur in our youth—so much so that Aristotle doubted that a person raised in vice could later change his character very significantly. Whether he is right about that is a matter primarily for psychologists, rather than philosophers. But it is interesting to note that on Aristotle's account there is a substantial element of moral luck² in the nature and quality of our lives. If we are fortunate enough to have been raised properly, so that we develop virtues, then our later lives will be much better for it. And those who were raised in immoral conditions, through no fault of their own, will likely turn out to be more vicious than virtuous.

The point of giving children a moral education is straightforward—to help them acquire the virtues. The key to this is developing their capacity for moral wisdom. Here, it can be instructive to think of children as apprentices being taught to gain a very complex skill, that of moral wisdom.

Think first about how apprentices in other areas are trained. An apprentice in a professional kitchen begins with a list of dos and don'ts, a set of hard-and-fast rules. Over time, she learns the limits of these rules, when to honor and when to break them. There is no master rule book that can give her this knowledge. She acquires it through trial and error, through the advice of experts, through a deeper understanding of cooking methods and of her ingredients. By the end of a successful education, she is something of an artist.

Aristotle himself drew an analogy with a navigator. Those just learning how to steer a ship would need to rely on firm rules to know how to get

a boat from one point to another. But with time and experience, an expert navigator will know when to bend those rules, or discard them entirely.

The same holds true of moral education. We begin as apprentices, following in an unquestioning way the rules handed down by our parents and teachers. In the early stages of their moral training, children learn simple rules, and are told to treat them as absolute: never, ever lie, steal, hit others, tattle, etc. These rules are crude, but it's right to ask our children to obey them. We address our learners where they are.

As children mature, they will, through experience and guidance, come to appreciate when exceptions are called for. We gradually step back from the rules we learned on our mother's knee, and subject them to careful scrutiny. A successful education will produce an independent thinker, one who doesn't need the old, over-simple rules as a crutch to get through each new situation. We understand, for instance, that honesty is the best policy. But sometimes honesty would be so hurtful and gain so little that evasion is the right way to go. As a rule, friends deserve our loyalty. But that doesn't mean that we must cover up for them if they steal from their employer and ask us to lie about it.

This line of thought supports the virtue ethicist's rejection of a simple moral litmus test, a formula that could be used by anyone no matter her degree of moral sophistication. Such a test not only overlooks the great complexity of morality, but also ignores the point that people possess moral wisdom in degrees. Advice that is suitable for a novice will be too crude for an expert, and vice versa.

The Nature of Virtue

The ultimate goal of a moral education is make ourselves better people. A better person is a more virtuous person—someone who is more courageous, just, temperate, and wise (among other things).

Virtuous people know what to do, when to do it, and why their actions are important. They have practical wisdom about things such as facing danger (courage) and giving people their due (justice). Indeed, practical wisdom is at the heart of every virtue. Virtuous people have insight and understanding about what each new situation requires of them.

A virtue is a character trait. It's not a mere habit, or a tendency to act in certain ways. Habits don't define a person; character traits do. The man who never shoplifts, despite feeling a strong temptation to do just that, has only the outward appearance of virtue, and not virtue itself. We can't

know that people are virtuous just by noting their actions, since virtues are defined not only by a pattern of behavior, but also by one's perceptions, thoughts, and motives. We need to know about a person's inner life before we can credit her with being virtuous.

Some people are loyal or generous by habit. Yet they may lack understanding about when and why it is appropriate to act this way. People might by nature be slow to anger, but anger is sometimes called for, and only the virtuous person will know when that is. Even if our nature inclines us to virtue, it can't get us all the way there. Virtue requires experience, practice, and reflection. The practical wisdom and the ethical motivations at the heart of every virtue distinguish it from a mere tendency to act in various ways.

To have a virtue is to see things from a particular perspective, to want certain things and not others, to rule out various options, to act in specific ways, and to have certain emotions and feelings. A virtue expresses a commitment to various ethical values, a commitment in thought, motivation, and deed. Virtuous people know what is right and why it is right, and want to do things because they are right.

Courage, for instance, requires that we correctly perceive various threats or dangers, control our fear in a reasonable way, be moved by a noble end, and act accordingly. Though Aristotle considered courage primarily in the context of the battlefield, this virtue, like all moral virtues, has its place in any number of more ordinary situations. The new kid in school displays courage when taking an unpopular stand among those whose approval and companionship he hopes for. Gandhi displayed courage in peacefully resisting the night sticks and attack dogs of the British colonial police. A whistle-blower may exhibit courage when revealing the corruption of her employers, knowing that she may be fired or sued for telling the truth.

People are virtuous only when their understanding and their emotions are well integrated. A virtuous person who understands the right thing to do will also be strongly motivated to do it, without regret or reluctance, for all the right reasons. In Aristotle's book, and in the virtue ethical tradition, this is what distinguishes the truly virtuous from the merely **continent**—those who can keep it together, manage to do the right thing, but with little or no pleasure, and only by suppressing very strong contrary desires. As Aristotle insists, "Virtuous conduct gives pleasure to the lover of virtue."³ This is one way to distinguish the truly virtuous from the merely continent.

Virtue and the Good Life

Aristotle thought it obvious that all of us seek *eudaimonia*, which translates as “happiness,” or “flourishing.” A life of *eudaimonia* is an excellent life for the person living it. The happiness Aristotle speaks of is not mere enjoyment. It isn’t only a state of mind, but rather a combination of activity and pleasure. Aristotle thought that the good life is an active one filled with wise choices and worthy pursuits. No matter how much pleasure you get from sitting in front of the TV and watching *The Simpsons* (a lot, in my case), a life devoted to that, no matter how pleasant, fails to qualify as a good life. Aristotle was no hedonist.

Aristotle argued that virtue is an essential element in a good life. In this he agreed with his teacher, Plato. Unlike Plato, however, Aristotle did not think that we could be happy on the rack. Virtue does not guarantee a good life; it is necessary, but not sufficient, for our flourishing. Most contemporary virtue ethicists side with Aristotle on this one—enough misfortune can damage a life so greatly as to make it, on the whole, an unenviable one. If a virtuous person loses his beloved family to war or disease, or falls prey to nasty rumors, crushing debt, and crippling disability, then no matter how virtuous, he can fail to gain true happiness.

But is virtue really essential to a good life?⁴ What about all of those criminals who get away with their crimes and enjoy a lavish retirement? What of the powerful tyrant who dies at the end of a long reign, peacefully and in his sleep? Few would argue that such people are virtuous. Yet they seem to live very good lives.

Virtue ethicists argue that appearances here are deceiving. Such people may indeed be pleased with how things are going, and get a lot of enjoyment from their lives. (Then again, a closer look at the criminal’s fear and insecurity, his emotional immaturity and the complications this brings, may make us think twice.) But virtue ethicists deny that pleasure is the be-all and end-all of a good life. Pleasure without virtue is not worth much. The most pleasant life may be a poor one indeed.

How can such a view be defended? Aristotle set up a three-part test to determine our ultimate good, in part to show that pleasure, wealth, power, and fame are *not* what life is all about.

First off, our ultimate good must not be something that is only instrumentally valuable. This explains why money and fame are ruled out, since these things have no worth of their own. They are merely a means to gaining other things of value.

Our ultimate good must also be *self-sufficient*. Possessing it is, all by itself, enough to make a life a worthy one. Political power fails this test. Having power over others is not what makes life valuable. Power is important, when it is, just because of what it enables a leader to do.

Finally, our ultimate good must involve something that is *distinctive* about us, something that is uniquely human. We need food to survive. But being nourished cannot be our final good, since we share this need with plants and animals. And since animals can experience pleasure, the point of our lives cannot be to gain pleasure, either. What sets us apart from everything else in the world is our rationality. Our ultimate good, then, must take the form of exercising our rationality. But there is little good in reasoning poorly. Rather, our ultimate good consists in the excellent use of our reasoning powers. And that is precisely what the virtues involve.

How attractive is a life of virtue? Very. Just think of what we hope and try for when raising our children. We want them to be kind, fair, generous, appropriately self-confident, and wise. We hope that they develop courage, that they know how to be a good friend, that they can sensitively offer comfort to others in need. Each of these is a virtue; a person who manages to have them all is in most ways living an excellent life. True, if Aristotle is correct, having these traits will not guarantee a good life. The admirable nature of a virtuous person may, for instance, attract the envy and hatred of others, who will sometimes make a martyr of a noble soul. But this should not lead us to think that the virtues are unnecessary for a good life. Even if a life of virtue is not a guarantee that you will flourish, a life without virtue is a poor one.

The virtue ethicist thus has an answer to a skeptic who charges the good person with being a dupe, with sacrificing self-interest on the altar of virtue. Being virtuous will (barring disaster) make you better off. It will ensure that you aim at things worth trying for. Virtuous people ordinarily do very well for themselves, even if the vicious sometimes have more fun. That is because human well-being is *defined* in terms of the virtues. Virtues are those excellences of character that contribute to one's well-being. Without them, one is leading the life of an animal—or worse.

Objections

The virtue ethical approach to life has a number of attractive features. I've tried to sketch some of the more important of them here. But given its unorthodox approach to morality, it is hardly surprising that virtue ethics has come in for its share of criticisms. Here are some of the more significant ones.

Tragic Dilemmas

Consider two central claims of the virtue ethical approach to morality:

1. Actions that would be done by a virtuous agent, acting in character, are morally right.
2. Such actions, when motivated by virtue, deserve our praise.

If these views are problematic, then virtue ethics is in deep trouble. Tragic dilemmas highlight the difficulty here.

A tragic dilemma is a situation in which a good person's life will be ruined, no matter what she does. All of her options will lead to disaster. Virtuous people will usually be able to avoid these situations, since they typically arise as a result of some serious moral mistake. Think, for instance, of the premise of so many movies—simple-minded guy finds bundle of cash, stupidly walks away with it rather than reporting it, and eventually faces a host of deadly choices.

But it is possible to find yourself in a tragic dilemma through no fault of your own. Consider the title character in William Styron's *Sophie's Choice*, who is detained in a concentration camp and then given the terrible news: one of her two children will be sent to the gas chamber. She must choose which one. If she refuses, both children will be killed.

Sophie's life will be ruined no matter what she does. But she should not withdraw from the situation; she must make a choice. A virtuous person (acting in character) would do so, since that would mean saving one of her precious children. If virtue ethics is correct, then selecting one of her children to be murdered is morally right and morally praiseworthy. But that seems wrong.

Here is an *Argument from Tragic Dilemmas* designed to set out the worry:

1. If virtue ethics is the correct account of morality, then Sophie's selection of one of her children to be murdered is morally right and morally praiseworthy.
2. It is neither.
3. Therefore, virtue ethics is not the correct account of morality.

Premise 1 assumes that a virtuous person would do as Sophie did—namely, select one of her children to die. Further, it assumes that there is a right way and a wrong way to do this. A virtuous person will see the tragedy for what it is, will not leap at the chance to make the selection, will not express joy at what is about to happen. And that was Sophie's response. She

was motivated as a virtuous person would be motivated—with a recognition of the terrible burden she faced, with love for her children, and with the greatest possible regret.

The only way to challenge premise 1 is to deny that a virtuous person in Sophie's shoes would select one of her children to be killed. That *might* be right—a virtuous person might refuse to make a deal with evil, and so try to keep her hands clean by not playing along with the sadistic choice offered to her. But recall that this means the death of both of her children, and it is hard to think that a virtuous person would prefer that to the death of one. I might be wrong about this. If so, then we have a way of rejecting the first premise.

That said, I think that the better option for the virtue ethicist is to criticize premise 2. *Under the circumstances*, a virtuous person *would* try to minimize the number of innocent deaths. And that means having to make a tragic choice, rather than refusing to do so. Choosing is indeed the right thing to do—even if it is absolutely heartbreaking.

It may also be praiseworthy. We might say of a person who refused to make this choice that, however understandable it may be, she was still being squeamish, and showed a lack of nerve. Having to make such a choice under these circumstances requires courage and fortitude, which is praiseworthy. Sometimes life presents us only with a choice among evils. Finding the inner strength to choose the lesser evil on that occasion need not be a moral failing, but may instead be something quite admirable. Our admiration should only increase when the choice involves an outcome that predictably destroys all of one's hopes for happiness.

If this analysis is correct, then virtue ethicists have an adequate reply to the Argument from Tragic Dilemmas. They can argue that certain choices in these situations are virtuous and that they are therefore right and admirable, even if, in more ordinary circumstances, any such choice would be purely evil.

Does Virtue Ethics Offer Adequate Moral Guidance?

Critics of virtue ethics often accuse it of failing to provide enough help in solving moral puzzles. When we are deliberating about how to behave, we'd like to have something more than this advice: do what a virtuous person would do.

But virtue ethics *can* provide more advice. It will tell us to act according to a large number of moral rules, each based on doing what is virtuous or avoiding what is vicious: Do what is temperate, loyal, modest, generous, compassionate, courageous, etc. Avoid acting in a manner that is greedy,

deceitful, malicious, unfair, short-tempered, etc. The list of virtues and vices is a long one, and this may really be of some help in figuring out what to do.

Still, the virtue ethicist has to face the familiar problem of moral conflict. What happens when these virtue rules conflict with one another? Suppose, for instance, that you are on vacation and happen to see your best friend's husband intimately cozying up to another woman. Would a virtuous person reveal what she has seen? Well, there is a virtue of honesty, and that points to telling your friend. But being a busy-body and rushing to judgment are vices; it's their marriage, not yours, and poking your nose into other people's business isn't a morally attractive thing to do.

All well and good. But you must do something. How to resolve this conflict (and countless others)? There *is* a right answer here, because there is something that a virtuous person would do. But virtue ethicists have offered very little instruction for deciding what that is. Once you appreciate which virtues and vices are involved in the situation, it is up to you to sort out how to balance them against one another.

This, of course, will be deeply unsatisfying to many people. They want their ethical theory to provide a clear rule that can tell them exactly what is required for each new situation. With expectations set this high, virtue ethics is bound to disappoint.

Unsurprisingly, however, virtue ethicists think that such expectations are implausible and far too demanding. They deny that ethics is meant to provide us with a precise rule or mechanical decision procedure that can crank out the right answer for each morally complex case. Recall the virtue ethicists' earlier criticisms of such an idea, and their claim that moral advice must be offered based on a person's level of wisdom and experience. There is no uniform moral guidebook, no formula or master rule that can tell us how to behave. We must figure it out for ourselves, through reflection, discussion, and experience.

Virtue ethicists can also argue that their theoretical competitors face similar problems. Most ethical theories incorporate a rule requiring promise-keeping. But isn't it sometimes okay to break this rule? If so, is there any *other* rule that could tell us precisely when we may break our promises? Try it out: You are allowed to break a promise if and only if _____. I don't know how to fill in that blank. That of course doesn't show that it can't be done. But anyone who can do it will also be able to know, in difficult situations, how to balance the virtue of fidelity against other considerations.

The bottom line is that almost every moral theory will require us to exercise good judgment in applying its rules. Virtue ethics requires more of us in this regard than some other theories, but that is a drawback only if morality can be made more precise than virtue ethicists believe. Whether that is so remains to be seen.

Is Virtue Ethics Too Demanding?

Virtue ethics tells us to do what a virtuous person would do in our situation. But what if a truly virtuous person sets a standard of excellence that is (almost) impossible to reach?

In 1933, Mohandas Gandhi went on a hunger strike that nearly killed him. Others have protested injustice by fasting unto death. Some of these protests were not based on personal grievances, but were expressions of outrage at social injustice. Assuming that some hunger strikers are virtuous people, acting in character, it appears that virtue ethics requires us to follow their lead.

Morality can sometimes require a great deal of us, but this may be going too far. One possibility, of course, is that it isn't. Perhaps we should be much readier than we are to give up our health or even our lives in political protest. Virtue ethicists could argue, as consequentialists have long done,⁵ that morality really does demand much more of us than we think. They might say that the expectations we've been raised with are too lax. If we were raised in a way that repeatedly reminded us of the importance of noble sacrifice, then we would be much more inclined to follow such examples. Our reluctance to sacrifice ourselves is no strike against virtue ethics, but rather against our own self-indulgence and desire for comfort and security.

Virtue ethicists could take a less severe stance, however, and argue that such extreme measures are appropriate only in rather special circumstances. If Russ Shafer-Landau went on a hunger strike, few would pay any attention, and so my extended fast would likely do more harm than good. The test of right action is to ask how a virtuous person, *in my circumstances*, would act. Since my circumstances are quite different from those of a world-renowned political leader, it doesn't follow that a hunger strike is something I should try myself.

That doesn't quite let me off the hook. For a truly virtuous person might do much more for others, and far less for himself, than I typically do in my everyday existence. And were he in my shoes, this might *still* be the case. So virtue ethics may indeed demand quite a lot from us.

Who Are the Moral Role Models?

If virtue ethics is correct, then we can resolve moral puzzles only by knowing how a virtuous person would act in our situation. Yet who are the moral exemplars? How do we decide who our role models should be, especially if different people endorse different candidates?

This is a very hard problem. For we pick our role models in large part by seeing how well they live up to our preexisting assumptions about what is right and wrong. Some people exalt suicide bombers as role models; others get sick just knowing that's so.

One solution to this problem is relativism—the idea that appropriate role models will differ from person to person, or culture to culture. This leads to the view that moral standards, too, will differ in this way. Since we spend a good deal of time on relativism in chapter 19, I suggest we move on and consider some alternative solutions here.

People can be truly virtuous even if we don't realize that they are. When we fail to land on the right role models, this is often explained by our own failure of virtue. For instance, it may be impossible to convince a cynical thief of the goodness of someone who conscientiously pays her taxes and keeps her promises. Winston Churchill, possessed of a great many virtues himself, nevertheless had racist attitudes toward Indians and was prepared to let Gandhi die in one of his hunger strikes, so committed was he to maintaining British rule over India. Churchill could not see Gandhi as a moral exemplar, but only as an annoying thorn in his side.

We become more insightful in selecting moral exemplars only by becoming morally wiser in general. And as we have seen, there is no fixed recipe for doing this. Moral education is a lifelong affair, and we are never fully wise. So we may indeed be off target in selecting our role models.

This isn't the whole story, of course. The whole story would involve a much more detailed account of how we gain moral knowledge, including knowledge of how to correctly identify our role models and how to resolve disputes about this matter. But in this respect, the virtue ethicist is in the same boat as everyone else. *Every* moral theorist has to answer hard problems about how we gain moral wisdom, and how to resolve disagreements about fundamental moral issues.

Conflict and Contradiction

We have seen in previous chapters how certain kinds of moral conflict can yield contradiction.⁶ Contradictions are a fatal flaw in any theory.

Virtue ethics may be saddled with contradictions, and if that is so, then it is sunk.

The problem is simple. If there are many virtuous people, then what happens if they disagree about what to do in a given situation? If, in my shoes, some good people would act one way, and others would behave differently, then it seems that the same action would be both right (because some role models would do it) and not right (because others would refrain from doing it). This is a contradiction.

The very wise people I have known do not all think alike. They don't see every case in the same light. They temper justice with mercy to varying degrees. They disagree about the role and form that discipline should take in good parenting. Some are more optimistic than others; some are more willing to demand more personal sacrifice than others. It thus seems possible that virtuous role models, acting in character, would do different things in the same situation. And that would yield contradiction.

There are a few ways out of this problem.⁷ The first is to insist that there is really only a single truly virtuous person, and so the differences that cause the contradictions would disappear. The second is to insist that every virtuous person, acting in character, would do the same thing in every situation. I don't find either of these replies very plausible, but perhaps there is more to be said for them than I am imagining.

The better option, I think, is to slightly modify the virtue ethical view of right action, given earlier in this chapter (p. 241) by the thesis labeled (VE). Assuming that virtuous people, acting in character, will sometimes do different things in the same situation, we should say that

1. An act in a given situation is morally required just because *all* virtuous people, acting in character, would perform it.
2. An act in a given situation is morally permitted just because *some but not all* virtuous people, acting in character, would perform it.
3. An act in a given situation is morally forbidden just because *no* virtuous person would perform it.

This really will solve the contradiction problem. If different virtuous people would act differently in the same situation, then we are no longer forced to say that an act is both right and wrong. Rather, we say that it is simply permitted, neither required nor forbidden.

A new problem arises, however. For if different virtuous people would act differently were they in our shoes, then how are we to decide for ourselves whom to follow? Which of the competing role models should we look to?

This isn't the same problem as the one discussed in the previous section (*Who Are the Moral Role Models?*). There, we were concerned with distinguishing between the virtuous and the not so virtuous, given that there is disagreement over who belongs in each camp. Here, we agree on who counts as a good role model. The problem is that there is more than one, and when they differ in their actions or advice, the theory doesn't tell us whom to follow.

This is another way in which virtue ethics fails to give us all the advice we might hope for from a moral theory. This isn't by any means a knock-down criticism of the theory. But it does point to more work that the virtue ethicist needs to do.

The Priority Problem

How do we get a handle on the nature of virtue? Here is the standard way. We first get clear about our duty, and then define a virtue as a character trait that reliably moves us to do our duty for the right reasons. So, for instance, to understand the nature of the virtue of generosity, we first determine that giving to others in need is right, and then define generosity as the character trait of giving to others in need, for the right reasons.

Virtue ethicists reject this strategy, because they deny that we can know our duty before knowing how virtuous people characteristically behave. For them, virtue has a kind of priority over duty—we must know what virtue is, and how the virtuous would behave, before knowing what we must do. Virtue ethics is unique in this regard. All other moral theories think of duty as the primary moral concept. For them, we can understand virtue only after we have the concept of duty under our belt.

The issue is about which concept is morally fundamental—virtue, or right action. To help see the stakes here, consider this question: Are people virtuous because they perform right actions, or are actions right because virtuous people perform them? Other moral theories go with the first option. Virtue ethics takes the second. And this raises a number of concerns.

Consider the evil of rape. The virtue ethicist explains its wrongness by claiming that virtuous people would never rape other people. But that seems backward. It is true, of course, that virtuous people are not rapists. But their rejection of rape is not what explains its wrongness. Rape is wrong because it expresses contempt for the victim, sends a false message of the rapist's superiority, violates the victim's rights, and imposes terrible harm without consent. We explain why virtuous people don't rape others by showing why rape is wrong. We don't explain why rape is wrong by showing that good people will not rape others.

The same goes for right actions. A bystander who sees a pedestrian about to fall into an open manhole should yell out and rush over to prevent the accident. Why? Not because a virtuous person would do such a thing (though of course she would). The real reason is to save an innocent person's life, or at least to prevent her from being seriously injured. It's not that intervention is right because virtuous people would do it; rather, they would do it because it is right.

If this has a familiar ring to it, that's because the structure of this theory closely mirrors that of Euthyphro's preferred view, the divine command theory (discussed in chapter 5). That theory denied that we could understand our duty apart from the decisions made by God, because God's commands are what create our duty. Virtue ethics takes a similar approach to morality, though many of its versions, including Aristotle's, are secular.

Virtue ethics tells us that it is the actions of virtuous people, rather than God, that determine what is right or wrong. People aren't virtuous because they do right; actions are right because they are done by the virtuous.

Virtue ethics and the divine command theory share a basic structure. And they share a basic weakness. We can see this by posing a familiar dilemma. Virtuous people either have, or don't have, good reasons for their actions. (1) If they lack good reasons, then their actions are arbitrary, and can't possibly serve as the standard of morality. (2) If they do have good reasons to support their actions, then these reasons, and not the actions themselves, determine what is right and wrong.

The second option is the better one. We must suppose that virtuous people act on good reasons, or else they wouldn't really be virtuous. Consider again the immorality of rape, and the many reasons why it is wrong. A virtuous person is one who is aware of these reasons and takes them to heart. Rape is not wrong because good people oppose it. They oppose it because it is wrong.

This approach preserves the integrity, the wisdom, and the goodness of the virtuous person. But there is naturally a cost. And it is steep. The cost is that the virtue ethicist's account of right action is directly threatened. That account tells us that acts are morally right *just because* all virtuous people would perform them in the circumstances, and wrong just because such people would refrain. But as we have seen, the choices of virtuous people do not make actions right or wrong.

We can still look to virtuous role models for reliable guidance on how to act. But their choices do not turn otherwise neutral actions into ones that are right (or wrong). They are not so powerful as that. Virtuous people

have keen insight into the reasons that make actions moral or immoral. They feel the compelling force of these reasons, and act accordingly. That is what makes them virtuous.

If this line of criticism is on target, then we have an explanation of why so many moral theories give priority to duty over virtue. We need to explain virtue in terms of duty, because we would otherwise be left with a picture of virtuous people that makes their choices arbitrary. But if that is so, then virtue ethics is in trouble, since one of its fundamental points is that rightness is defined in terms of the choices of the virtuous.

Conclusion

Virtue ethics represents an exciting continuation of an ancient tradition. It has a variety of attractions, not least of which is its emphasis on the importance of moral character. It represents a pluralistic approach to morality, and has interesting things to say about ethical complexity, moral education, the importance of moral wisdom, and the nature of the good life. Many of the criticisms that have been leveled at it can be met once we dig a bit deeper, or introduce small changes to the theory.

But no ethical theory, at least in its present state, is immune to all real difficulties, and virtue ethics, too, has its vulnerable points. The greatest of these takes aim at one of its central claims: that right action must be understood by reference to virtue, rather than the other way around. Perhaps virtue can really enjoy this sort of priority. But it will take a great deal of further work to show it so.

Notes

1. See the following discussions for more in-depth treatment of why people would have such hopes: on the structure of moral theories, pp. 12–14; on proceduralism, pp. 177–178; on knowing one's duty, pp. 227–234.
2. For more on moral luck, please see chapter 12, pp. 172–173.
3. *Nicomachean Ethics* 1099a12.
4. For more on this topic, see chapter 8, pp. 102–105, and chapter 14, pp. 191–195, as well as the general discussion in part 1.
5. See chapter 10, pp. 130–133, for more discussion of how demanding consequentialism can be.
6. See especially chapter 15, pp. 212–213. This matter is also discussed at some length in chapter 19, pp. 284–288.
7. A similar problem confronts the social contract theory; see pp. 198–199.

..........
Feminist Ethics

The Elements of Feminist Ethics

The most prominent authors and supporters of the ethical theories that we have considered so far have one thing in common. They are all men. Most of them lived in societies that systematically discriminated against women. Since even the most high-minded philosophers are bound to reflect some of the common assumptions of their times, it should come as no surprise that many philosophical luminaries held views about women that nowadays make us cringe.

Aristotle said that “the male is by nature superior, and the female inferior; the one rules, and the other is ruled.”¹ Aquinas claimed that “As regards her individual nature, each woman is defective and misbegotten.”² Kant wrote that “laborious learning or painful pondering, even if a woman should greatly succeed in it, destroy the merits that are proper to her sex...they will weaken the charms with which she exercises her great power over the other sex... Her philosophy is not to reason, but to sense.”³ Rousseau said that “Women do wrong to complain of the inequality of man-made laws; this inequality is not of man’s making, or at any rate it is not the result of mere prejudice, but of reason... [Women] must be trained to bear the yoke from the first, so that they may not feel it, to master their own caprices and to submit themselves to the will of others.”⁴

We might be tempted to downplay these slights by claiming that they did not influence the main lines of argument of these thinkers. And there is a sense in which this is correct—almost none of the major

male philosophers of past centuries wrote very much about women. But there is also a sense in which it is incorrect. For there are two ways in which philosophers have shortchanged the lives of women. The first is to make false and damaging claims about them. The second is to ignore female experiences and perspectives. Both have been the norm in ethical thinking for centuries. **Feminist ethics** seeks to remedy both of these flaws.

Feminist ethics is not a single theory, but rather a general approach to ethics that is defined by four central claims:

1. Women are the moral equals of men; views that justify the subordination of women or downplay their interests are thus mistaken on that account.
2. The experiences of women deserve our respect and are vital to a full and accurate understanding of morality. To the extent that philosophers ignore such experiences, their theories are bound to be incomplete, and likely to be biased and inaccurate.
3. Traits that have traditionally been associated with women—empathy, sympathy, caring, altruism, mercy, compassion—are at least as morally important as traditionally masculine traits, such as competitiveness, independence, demanding one's fair share, a readiness to resort to violence, and the insistence on personal honor.
4. Traditionally feminine ways of moral reasoning, ones that emphasize cooperation, flexibility, openness to competing ideas, and a connectedness to family and friends, are often superior to traditionally masculine ways of reasoning that emphasize impartiality, abstraction, and strict adherence to rules.

Two cautionary notes. First, no one believes that every woman is compassionate and caring, or that every man is aggressive and competitive. To the extent that these generalizations hold true, they do so only for the most part. Second, when I speak of *traditionally* masculine and feminine traits, I mean just that. These are features that our cultures have long associated with men and with women, respectively. But there is no claim that such traits are innate. We realize now, as so many have failed to do in past, that many characteristics we associate with certain groups are a by-product of social influences. Stereotypes often fail to have any basis in fact. But even when they do, these facts are often a result of difficult circumstances and limited opportunities, rather than the expression of some inborn character.

Moral Development

Feminist ethics really came into its own in the 1980s. Before that, scattered writings by feminist philosophers had been appearing for at least two centuries. But in 1982, Harvard psychologist Carol Gilligan published *In a Different Voice*.⁵ It's fair to say that this book did more than any other to launch this new movement in philosophical thinking.

Gilligan argued that women think and experience the world differently from men. This was not news—psychologists had long agreed on this point. The difference, though, was that Gilligan rejected the mainstream views that saw women's thinking as inferior to men's.

One of the most influential models of moral thinking in the 1970s was put forth by Lawrence Kohlberg, a teacher and colleague of Gilligan's. Kohlberg defended the idea that there are six stages of moral development. At the earliest stage, children see moral rules only as potential threats, and they behave well only out of fear of punishment. As we grow, we view morality (at Kohlberg's third stage) as depending on our social roles and on our relations. Moral demands come from these roles and relationships; the point of morality is to reinforce them. Ultimately, at the sixth stage, we view morality as requiring obedience to abstract rules of impartial justice. They require us to see our situation dispassionately. They assign everyone equal importance. These rules are universal, and do not depend on, or refer to, the particulars of our character or our situation.

Gilligan noted that many women fare poorly on Kohlberg's scale—they never advance beyond the third stage. Their moral thinking takes very seriously the specifics of personal relationships, and only rarely involves appeal to abstract, universal principles to solve moral problems. Gilligan argued that this was not evidence of the moral immaturity of these women. Rather, it indicated the failure of Kohlberg's six-stage model.

Kohlberg was assuming that the moral ideal is someone who appeals to principles of justice that apply to everyone, regardless of social, religious, or political position. Gilligan did not downplay the importance of such ideals, but rather argued for the equal importance of a second model of moral thinking, one that placed an emphasis on seeking compromise, seeing the value in positions that conflicted with one another, and doing what was needed to maintain or repair personal relationships. In the studies that formed the basis of her book, Gilligan repeatedly found that women tend to follow the second model. It was only a prejudice—a very

well-entrenched and long-standing one, to be sure—that placed the first model above the second.

Gilligan did not argue that *all* women thought in these ways. Nor did she claim that women, *by nature*, were attracted to these ways of thinking. Still, she claimed to have identified a very strong tendency among the women she studied to react to cases of moral conflict in the same way—a way very different from that of the men she interviewed. The women brought an attitude of care and sympathy to their decision-making, an attention to the concrete particulars of the cases they were confronted with. They rarely appealed to abstract moral principles. They did not regard justice as all-important. They were partial to certain people—family members, friends, lovers—rather than seeking to give everyone the same degree of concern. In deciding what to do, they sought compromise where possible, and did not regard the relevant moral rules as absolute. They tried to “split the difference” in cases of conflict, and showed greater respect for views that differed from their own. They were inclined to voice their views with hesitation and humility, rather than with great confidence and assurance.

Gilligan is a psychologist, not a moral philosopher. She saw her work as describing the differences in male and female thinking, and did not focus on drawing out the ethical implications of her research. But philosophers soon took up this task, developing a new paradigm of moral thinking that set out to challenge the prevailing ethical wisdom on a number of fronts. Let’s turn our attention to some of the most important of these challenges.

Women’s Experience

There is always a danger of oversimplifying things when discussing “the female perspective.” There is no such thing, strictly speaking. Women’s experiences, their outlooks on life and their responses to it, are extraordinarily diverse. On any issue of importance, you will find women on opposite sides of the fence—just as we do men. Many men share the outlook that Gilligan associated with women, and many women share a more “masculine” approach to moral thinking.

And yet...even though there is no one thing that is “the female outlook,” there are some distinctively female experiences (giving birth and mothering, most obviously), and many others that are widely shared by women, less so by men, and that have been largely neglected by philosophers. One of these is a vulnerability to rape. Another is the threat of domestic abuse. Yet another is the systematic disparity in pay for equal

work. Another is women's total or near total exclusion from a variety of professions (president, airline pilot, welder, mason, fire fighter, electrician, etc.). This list could be expanded, surely. And it could go on for pages if we broadened the scope of our investigation, to consider the status of women in other countries that have done less to reduce overt sexual discrimination.

These threats and limitations are not suffered only by women. Many men have been raped and beaten, have suffered discrimination of various forms. But the likelihood of such suffering is greatly increased if one is a woman. And yet prior to the early 1980s, philosophers had almost completely ignored these issues, devoting little if any thought to the ethical problems they raised. One goal of feminist philosophers is to put these issues on the map, and to get philosophers and policy makers to devote their time and attention to solving them.

There is another, related aspect of ordinary female experience that bears mentioning here: an increased dependence and diminished autonomy. Compared to men, women have almost always had far fewer choices open to them, and have enjoyed far less control over important aspects of their lives. In many countries, women still need the permission of either their husbands or a close male relative to travel outside of their region. Women have been, and in many places still are, unable to choose a spouse for themselves. Women are far likelier than men or boys to be sold into sexual slavery. Many kinds of jobs, all political offices, and a variety of social positions have been closed to women, just because they are women.

These are commonplace observations—no less true and no less worrisome for that. We can summarize them by noting that a central fact of most women's lives is their dependence, and the frequent demands to place their own interests on the back burner. Married women have traditionally had very little choice about where to live, for instance—if their husband's job required a family move, then that was that. It is still the case that the mother is usually the one to sacrifice job prospects if a couple decides that one of them must stay home to care for their child. Historically, and still in most areas around the world, women are dependent on men for their economic survival. They are vulnerable to physical abuse by husbands or male relatives. They are legally at the mercy of judicial systems ruled by men.

Dependence on others often requires us to rethink our moral assumptions. Consider a popular one: We should stand up for our rights and defend our honor against attacks. But if a woman has little education and less job training, does she walk out on an abusive husband who is paying

the bills? Does she confront him and go so far as to seek a divorce—knowing the terrible statistics of spousal murder, and not knowing where her next meal will come from?

When we rely on others for our income, a roof over our head, or physical security, we are vulnerable to their decisions and constrained in our own. That is the situation that most women and all children face. It is also the situation of many men. To the extent that philosophers hope to address the real-life situations of most people, they must follow the advice of feminist thinkers and pay more attention to the ways in which we are dependent on others.

The choices of most women (and most men, for that matter) reflect their role within a family—as a grown child of elderly parents, as a spouse of someone whose needs and wants will not always match your own, as a sibling whose brothers or sisters may sometimes require your immediate, significant assistance. We do not choose our parents or our siblings. Consent does not enter into it, and yet here we are, with duties to them whether we like it or not. What morality requires of us often depends on who we are related to. And this is usually something that we have no control over.

The importance of vulnerability, of not having control over important aspects of one's life, of dependence and connectedness to others, are all features of the moral life highlighted by feminist ethics. What we ought to do with our lives, from the big picture to our everyday choices, is often a matter of having to recognize the importance of those we care about. The interests and desires of our partners or spouses must be taken into account. Those of us with children know how needy they can be—even the least needy among them. True friends are sometimes very demanding. Fulfilling our duties across this range of relationships often means setting aside our own interests, or seeing our interests as crucially dependent on the interests of those we care for.

It's not that women alone are deeply connected in this way. *Most* of us are enmeshed in a network of personal relationships, many not of our own choosing. Feminists argue that this should force us to rethink the ideal of a person as a free agent, isolated, and wholly independent. Once we really appreciate how connected we are to others, the moral philosophies that are based on ideals of self-interest or full autonomy may become less appealing.

A way of reinforcing this point comes from feminists who emphasize the importance of caring in establishing and sustaining our vital relationships. Women have usually been cast in the role of nurturer—the one who

sensitively cares for others in need. Of course not all women are tender and nurturing. But it is important to understand how much of what goes on in society depends on domestic work being well done—children fed and well raised, house and household well maintained. Society is possible only if there are people devoted to caring for others. Without caretakers, children would die, the sick and the infirm would die, and many of the mentally disabled and mentally ill would die. Most of those whose work does not involve caring for others are able to keep their jobs only because there is someone back at home who is managing the household. That someone is usually a woman. Where it is not a wife or mother, it is usually a domestic employee or servant—and this, too, is almost always a woman.

Where standard ethical theories see morality as primarily about the pursuit of self-interest (egoism), doing justice (Kantianism), seeking mutual benefit (contractarianism) or impartial benevolence (utilitarianism), many feminists point to care—especially a mother's care—as the model of moral relations and the basis of ethics. This maternal model has generated what feminist philosophers now call an *ethic of care*. Let's take a closer look at this new moral theory.

The Ethics of Care

The major moral theories we have discussed thus far are not designed with home and family life in mind. But since so many of our most important moments are spent with those we love, and since so many moral choices are made within the context of close relationships, why not imagine what an ethic would look like that took these as its starting points? In particular, many feminist ethicists have argued that we should think of a loving mother's care for her children as a model for all moral behavior.

We can better understand an ethics of care by first seeing what it is not. Unlike ethical egoism, care ethics does not insist that we always look out for Number One. Mothers often rightly sacrifice their own interests in order to advance those of their children. Unlike Kantianism, an ethics of care does not place paramount importance on justice. Matters of justice are not entirely absent from parent-child relations, but they are certainly not the primary focus here. It is important that a parent not try to swindle her children, and that children show respect for their parents. But standing on one's rights, insisting on a fair share, and ensuring that the guilty are given their just deserts are not at the heart of loving relationships.

Contractarian theories see the authors of the moral law as indifferent to the needs of others, willing to make sacrifices for them only if there is a reasonable chance of being compensated in return. Good parents don't see things that way. A mother's care is not conditional on her child's obedience to a set of mutually beneficial rules. The rational pursuit of self-interest is not the ultimate goal; if the only way to help your child is to take a serious hit yourself, a good parent will often do just that.

And contrary to utilitarian demands for impartial benevolence, loving parents are much more concerned about their own children than about other people's kids. There is no thought of being impartial here; a good mother will be partial to her children, will give them more care and attention than she does anyone else's children. Love and care cannot be parceled out to everyone equally.

In addition to these specific differences, the ethics of care incorporates the following features. Most of these represent a point of departure from most traditional ethical theories, though as we'll see, there are some points of similarity between the ethics of care and both virtue ethics and Ross's ethical pluralism.

The Importance of Emotions

Care is an emotion, or a network of reinforcing emotions that involve some combination of sympathy, empathy, sensitivity, and love. Like all emotions, care has elements that involve thinking and feeling. The relevant thoughts are focused on the wants and needs of the one being cared for. The feelings are positive, friendly, helpful, nurturing, and often loving. Care helps us know what others need—mothers and fathers often understand what their own child needs much better than anyone else. And care helps to motivate us to tend to those needs, even when we are exhausted, begrudging, or angry. How many parents have roused themselves from a sound sleep to soothe their crying infant? Care helps ease those parents out from under the covers.

Utilitarians don't place much importance on the emotions in knowing what's right and wrong. Calculating amounts of happiness and misery isn't an emotional task. And we have seen how dismissive Kant was of the emotions, claiming that reason alone could both tell us where our duty lay and get us to do it.⁶ Kant was surely right in thinking that our emotions cannot go unchecked—we need an ethic of care, and not just care itself. But feminist philosophers argue that care and its associated emotions are

central to moral motivation and moral discovery, even if they are not the whole story.

Those who defend an ethics of care sometimes see themselves as working within a virtue ethics tradition. And this makes sense, given the emphasis not only on what we do, but on how we do it. The manner in which we do things is often as important as what we do. Suppose, for instance, that my mother calls me up and asks that I spend the afternoon helping my aged father with some household chores. I do as she asks, but only begrudgingly, and make it clear with my body language and my brusqueness that I resent being there. I've done the right thing, but in the wrong manner. I am not acting virtuously, and am not displaying an appropriate level of care.

Against Unification

Most of the traditional ethical theories offer us one supreme moral rule that determines the morality of all actions. Such a rule is meant to unify ethical theorizing by providing a single source for every moral duty. Feminist ethics rejects this picture. I can't offer a short formula to describe the feminist conditions of right action, because there is no such formula. Ethics cannot be systematized. If feminists are correct, there is no ultimate rule that can explain or justify all of our moral duties.

This has a number of important implications.⁷ One of these is that there is no formula or surefire litmus test for knowing what morality demands of us. Morality is complicated and messy. The drive to try to unify all of morality under a single supreme rule is an understandable one. Such a rule would lend clarity and structure to ethics. But feminists argue that this is a pipe dream.

We can see this as it plays out in the lives of many women (and men) faced with conflicting demands from children, work, spouses, and other sources. Suppose your parents call you up and proceed to criticize your boyfriend. He later asks you what you and your parents talked about. Do you tell him what they've said, knowing that he'll be hurt and that this is going to make a good relationship between him and your parents even harder to achieve? Or suppose your husband believes in disciplining children with a very firm hand. You disagree. He spanks his son—your son—after some minor misbehavior. Then he does it again. What do you do?

These aren't life-or-death cases. Rather, these are the sorts of everyday situations that arise in homes all the time. Feminist philosophers say of such examples that while there is often a right thing to do, we can't read

off a recommendation from some simple rule. Rather, we have to appreciate the different sources of our moral duties. These stem primarily from relationships we have with other people. And they can conflict with one another. When they do, it can be very hard to know what to do. At such times, we may wish for some easy formula that could give us instant advice about how to behave. But if feminist philosophers are right, there is no such thing. Part of gaining moral maturity is recognizing this, facing life's difficult choices, and not pretending that overly simple answers will solve our problems.

Against Impartiality and Abstraction

There are many reasons that philosophers have been so attracted to the idea of a supreme moral rule. Here is one of them. The more general and abstract the rule, the less likely it is to include bias. A rule that applies only to certain people or to certain situations may reflect only a limited perspective. Philosophers have long sought an outlook that is free of prejudice and distortion, one that takes into account all people at all times.

But why is this so important? The traditional answer is that it gives us a way to ensure impartiality. We must think of everyone as moral equals, and that means giving each person equal weight in determining what is right and wrong. But as we have seen, feminists reject the idea that we must proceed in this way. It is right that we give priority to those we care about. It is good to be partial to our loved ones.

Feminist ethicists resist the push to abstraction that we see so strongly in philosophy. Moral reasoning should not be centered around a single, very general rule, but rather should be guided by a more complicated understanding of the specifics of situations. For instance, feminists reject Rawls's attempt to strip away all concrete, particular knowledge of who we are when determining the principles of justice. (See p. 184.) They reject the utilitarian emphasis on impartiality. They deny that any general rule, such as *lex talionis* (see pp. 166–168), can illuminate the nature of justice. We must instead substitute a sensitive appreciation of the details and complexities of a situation, and not try to solve moral problems by seeing how these features fall under some general moral rule. If an ethic of care is correct, there is no master rule that can successfully guide our moral decisions.

Against Competition

It shouldn't be surprising that the virtues of competition are so often touted by men. Competition governs the business world, the realm of politics, and

the world of sports—for so long exclusively male domains, only recently opened to women. Moral theorists have sometimes borrowed this emphasis on the importance of competition. Ethical egoists require us to harm others if their interests compete with our own. Contractarians think of social interactions as a series of *prisoner's dilemmas* (see pp. 178–181), in which each person's interests are pitted against those of others.

An ethic of care seeks to replace this picture with one that values cooperation over competition. A healthy mother-child relationship is not a competitive one. It does not set the interests of parents against their children. It is marked by kindness and a willingness to sacrifice for one another. Good parents will see their interests as very closely bound up with the interests of their children, rather than in competition with them.

Feminists argue that we should try to turn competitive situations into cooperative ones. Rather than highlight the areas in which our interests clash, we should try to seek reconciliation. We should turn demands into requests, try to understand where the other person is coming from, and be flexible in our dealings with others. We can't show the same degree of care toward others as we do toward our loved ones, but we can approximate this, and look to care as the basis for restructuring our relationships. Imagine how different the business, political, or social worlds would be if people were animated by care rather than by self-interest and competitiveness. It's a far cry from the world we live in—and a far better one, if feminists are right.

Downplaying Rights

Feminists often argue that moral theories have placed too much emphasis on justice. Demanding our rights, insisting that others honor our claims, and making sure we get what we are entitled to—these are ways of asserting our independence from one another, rather than our connectedness. We stress fairness, rather than cooperation, when we are competing with one another. Talk of rights can divide us more quickly. This is a common complaint about the abortion debates, for instance. Once we start speaking of the rights of a fetus and of a woman, the debate becomes bogged down, making it very difficult to find common ground with those on the other side of the fence.

Imagine that we instead emphasized our responsibilities to one another, based on the model of a caring parent toward her children. Society would be seen not as a venue for the pursuit of rational self-interest, but rather as a stage for cooperation where we took responsibility for one

another, and especially for the most vulnerable among us. In the area of social policy, for instance, this would lead to placing much greater importance on education, support for poor families, and on ensuring that everyone had access to excellent medical care.

The emphasis on rights has often meant giving priority to our being free from coercion and unwanted interference. Rights protect autonomy and independence. And so we have rights, for instance, to say and to read what we want, or to do what we like within the privacy of our own homes.

But many (though not all) feminists have launched pointed criticisms of such priorities. They argue that rights tend to place us in opposition to others, creating a barrier beyond which no one may pass without permission. Individual rights often allow people to pursue their own paths at the expense of the community. Rights emphasize the ways in which we are separate from one another, rather than the ways in which we might be brought together. After all, loving parents do not stand on their rights when their child needs them. They do not want to assert their independence from their son or daughter. Feminists argue that rather than finding ways to insulate ourselves from others, we should be looking to create more opportunities for people to help one another. We should emphasize our responsibilities to others, rather than our rights against them. To the extent that rights stand in the way of building community and forging close ties with others, most feminists regard them with suspicion.

Challenges for Feminist Ethics

Feminist ethics is an approach to morality, rather than a single unified theory with specific claims that all feminists endorse. As a result, a presentation of this family of views must settle for highlighting the shifts in emphasis that feminists urge us to adopt, rather than in isolating particular arguments and views that all feminists will accept.

Feminist ethicists currently deal with several challenges. And this is unsurprising, given that extensive work in the area is only a generation old. Here are some of the most important of these challenges.

First, the feminist ethics of care threatens to restrict the scope of the moral community too greatly. Indeed, early care ethicists argued that we have moral duties only to those we care about. This view is no longer argued for, as it leaves us without any moral duties to strangers or to those we thoroughly dislike. But if we are to model our moral behavior on the

mother-child relation, then we need extensive advice about how this is supposed to work in the case of those we don't know or care about. After all, one way in which we seem to have made moral progress is by extending the scope of the moral community beyond those who are near and dear to us.

Second, the role of the emotions in helping us to know the right thing to do, and in moving us to do it, needs further exploration. Moral clarity sometimes requires that we overcome our indifference and become more emotionally invested in an issue. But in other cases, emotions can cloud our judgment. We need a view of which emotions are appropriate, and when they are appropriate, since the very same emotion can sometimes be enlightening and at other times, anything but. An emotion such as anger often blinds us to the truth and prevents us from doing right. And so it needs to be regulated. But anger can also correctly alert us to serious immorality, and will sometimes move us to overcome our fear and to do the right thing. We need a much fuller story about the role of the emotions in the moral life.

Third, downgrading impartiality has its costs. As we saw in an earlier chapter (9, p. 119), there is a great deal to be said for the importance of impartiality. It is a definite virtue of judges and others who hold positions of civic responsibility. It is an important corrective for prejudice and bias. It is one of the best reasons for taking the interests of women as seriously as those of men. Impartiality may not always be the right way to go, but it is, at least sometimes and perhaps usually, the best perspective from which to make important moral decisions.

Fourth, the rejection of a unified moral theory with a supreme moral rule leaves it unclear how we are to resolve cases of moral conflict. A virtue of the principle of universalizability, or the principle of utility, is that we have a definite standard to appeal to in trying to decide how to act in puzzling cases. Without such a standard, we may be left largely in the dark about what morality allows or requires of us.

Fifth, while cooperation is often an excellent thing, we also need to have strategies for dealing with uncooperative people or governments. The world would be a much better place if we were all able to get along and put our differences behind us. But as we all know, good faith and flexibility are sometimes met with a sneer and an iron fist, and we need contingency plans for such occasions. Caring for our enemies will sometimes mean that they kill us or those we are entrusted to protect. Further, competition is sometimes a good thing. It can enhance efficiency in business. It can make for inspiring athletic events. It can spur us to personal excellence. So

we shouldn't give up on competition entirely. And that means developing a sophisticated view of when it is and isn't appropriate to prefer cooperation over competition.

Sixth, while justice and rights are not the whole of morality, they are nonetheless a very important part of it. We can explain what is so immoral about the oppression of women by citing the rights that are violated by sexist actions and policies. Women have rights to be free of physical abuse; they have a moral right to be paid the same amount of money for doing the same work; it is a gross injustice to forcibly circumcise a teenage girl (or a grown woman, for that matter). Rights are a form of moral protection, and women are often the ones in need of the strongest protections. A plausible feminist ethic must therefore make room for the importance of moral rights and the demands of justice that they support.

Conclusion

Feminists have often been described as those who think that women ought to be treated exactly as we treat men. But this is a mistake. Feminists argue not for equal *treatment*—after all, many of the ways that men typically get treated are morally questionable. Rather, feminists argue for *equal consideration*. The interests of women are to be given the same importance as those of men. When setting social policies, when evaluating traditions, or when trying to settle conflicts between men and women, it is immoral to downgrade the interests of women just because they are women. Women are the moral equals of men. This simple idea, if taken seriously, would lead to radical change in most areas of the world.

Perhaps Carol Gilligan was right in arguing that women's moral thinking typically differs significantly from that of men. But even if she wasn't (as a number of psychologists have argued), we can still learn a great deal about ethics by reimagining moral philosophy through the lens of traditionally feminine concerns. Many of us, men as well as women, are more vulnerable and dependent than traditional moral theory allows. In the real world, there are severe inequalities of wealth and power, and it pays to be sensitive to such things when deciding on our moral ideals. Making care the centerpiece of our moral life, and allowing emotions and our loving relations a larger role in moral thinking, can make a substantial difference in our ethical outlooks.

Feminist ethics is not just for women. Its recommendations are intended for men and women alike. The importance of care, and emotions generally; the emphasis on cooperation; the attractions of flexibility and

compromise; the need for more than justice—each of these is as morally important for men as it is for women.

Feminist ethics is best seen as a general approach to morality, rather than as a well-developed theory that can at this point compete directly with the traditional moral theories. But this is not necessarily a weakness. Rather, it is evidence of the wide variety of views that can be developed by those who take the interests of women just as seriously as we have long taken those of men.

Notes

1. Aristotle, *Politics* 1254 b13.
2. Thomas Aquinas, *Summa Theologica*, Question 92, First article.
3. Immanuel Kant, *Observations on the Feeling of the Beautiful and the Sublime*, section 3.
4. Jean-Jacques Rousseau, *Emile: On Education*.
5. Harvard University Press, 1982.
6. See the discussion in chapter 12, pp. 164–165.
7. Most of these have been discussed in detail in chapters 15 and 16, which are devoted entirely to various forms of ethical pluralism. Pluralist theories defend the idea that there are a number of different fundamental moral duties, rather than just one. Virtue ethics also endorses this idea; see pp. 242–243 for more discussion.



PART THREE

The Status of
Morality



Ethical Relativism

Moral Skepticism

Each of us has our doubts about morality. Some of these doubts focus on its content—we aren't sure of what morality allows, and don't know where our duty lies. This is a common worry, absolutely familiar to most of us.

But there is another kind of doubt, of a less usual, but no less disturbing sort. Indeed, this kind of worry can undermine all of our confidence in morality. This sort of puzzlement is not about, say, whether lying is ever acceptable, or whether we may break a death-bed promise to a loved one. Rather, it is a “deeper” worry, one that puts aside all specific debates about what is right and wrong, and asks instead about the entire enterprise of morality. The worry, specifically, is that **moral skepticism**¹—the denial of objective moral standards—is correct, and that morality therefore lacks any real authority.

The notion of **objectivity**², like so many others that we have seen in these pages, is ambiguous. Objective moral standards are those that apply to everyone, even if people don't believe that they do, even if people are indifferent to them, and even if obeying them fails to satisfy anyone's desires. Moral claims are objectively true whenever they accurately tell us what these objective moral standards are, or what they require of us.

There are millions of objective truths. Here are three, at random: The planet Jupiter has a greater mass than Mercury. John Milton wrote *Paradise Lost*. Galileo is dead. It doesn't matter what you think of these claims, and it doesn't matter what I think of them. It doesn't matter whether I care about these claims, and it doesn't matter whether believing them satisfies

any of our desires. Neither personal opinion nor conventional wisdom makes these claims true. They are true and would continue to be so even if no one believed them.

But are there any objective *moral* truths? That's not so clear. There are plenty of reasons for doubt; the most popular and important of these will be the focus of our final chapter. If such doubts are correct, then **ethical objectivism** must be false. Ethical objectivism is the view that some moral standards are objectively correct and that some moral claims are objectively true.

Before having a look at these criticisms of ethical objectivism, let us consider the alternatives. This requires that we sort out the various forms that moral skepticism can take. (And it means just a little more jargon.) There are basically two forms of moral skepticism: **moral nihilism**, and **ethical relativism**.

Moral nihilism is the view that there are no moral truths at all. Taking a close, hard-nosed look at what is real and what isn't, nihilists place morality squarely in the latter camp. The world contains no moral features. Don't be fooled by our common talk of genocide's immorality or a murderer's evil nature. That sort of talk is either just plain false, or a disguised way of venting our feelings (of hatred, disgust, etc.)

According to the moral nihilist, when we take a step back from the issues that engage our emotions, we can see that nothing is right, and nothing wrong. The world will one day be fully described by science, and science has no need of moral categories. In the words of the brilliant Scottish philosopher David Hume (1711–1776), we *gild* and *stain* a value-free world with our feelings and desires. When we declare a murderer wicked or a relief worker good and kind, we are expressing our anger or our admiration. We are not stating a fact. We couldn't be, since there is no moral reality to describe. As a result, no moral claims are true.

By contrast, ethical relativists claim that some moral rules really are correct, and that these determine which moral claims are true and which false. Many are true. People sometimes get it right in ethics, and they do that when their beliefs agree with the correct moral standards. Crucially, these standards are *relative* to each person or each society. A moral standard is correct just because a person, or a society, is deeply committed to it. That means that the standards that are appropriate for some people may not be appropriate for others. There are no objective, universal moral principles that form an eternal blueprint to guide us through life. Morality

is a “human construct”—we make it up, and like the law, or like standards of taste, there is no uniquely correct set of rules to follow.

These two brands of moral skepticism are quite different from each other. It pays to treat them separately. We’ll keep the focus on ethical relativism in this chapter, and turn to moral nihilism in the next.

Two Kinds of Ethical Relativism

As you may already have noticed, ethical relativism isn’t just a single doctrine. It actually comes in two varieties: **cultural relativism** and **individual relativism** (usually referred to as **ethical subjectivism**, a name I’ll use from now on). Cultural relativism claims that the correct moral standards are relative to cultures, or societies; ethical subjectivism claims that the correct moral standards are those endorsed by each individual. The difference amounts to whether society, or each person, has the final say about what is right and wrong. This is undoubtedly an important difference, but as we’ll see, both the advantages and the drawbacks of cultural relativism and ethical subjectivism are remarkably similar.

Consider subjectivism first. It says that *an act is morally acceptable just because (a) I approve of it, or (b) my commitments allow it. An action is wrong just because (a) I disapprove of it, or (b) my commitments forbid it.* My commitments are the principles I support, the values I stand for. In this line of thinking, personal conviction is the ultimate measure of morality. Right and wrong are wholly in the eyes of the beholder.

Subjectivists think that there are right answers in ethics, but that these are always relative to each person’s moral standards. There is no superior moral code that can measure the accuracy of each person’s moral outlook. If subjectivism is correct, each person’s moral standards are equally plausible.

Cultural relativism instead locates the ultimate standard of morality within each culture’s commitments. It says that *an act is morally acceptable just because it is allowed by the guiding ideals of the society in which it is performed, and immoral just because it is forbidden by those ideals.*

Both subjectivists and relativists regard people as the authors of morality. In both of these views, morality is made by and for human beings. Before we were around, nothing was right and wrong. If our species ever becomes extinct, morality will cease to exist. The fundamental difference between these two views is whether each person, or each society, gets to have the final say in ethics.

Despite their disagreement about whose views are morally authoritative, both ethical subjectivism and cultural relativism share a number of similar elements that make it easy to evaluate them in tandem. Let's now have a look at some of the most important features of these views.

Some Implications of Ethical Subjectivism and Cultural Relativism

Moral Infallibility

Subjectivism and relativism occupy a middle ground between moral nihilism and ethical objectivism. There are legitimate moral standards (contrary to nihilism), but their legitimacy depends crucially on our support (contrary to objectivism).

But subjectivists and relativists do not always see eye to eye. Subjectivists are suspicious of cultural relativism because of their belief that societies can be deeply mistaken about what is right and wrong. If a social code can contain some serious moral errors, then cultural relativism is in trouble, since it says that whatever society holds most dear is morally right.

Relativists admit that some social beliefs can be morally mistaken. These are the ones that clash with society's most cherished ideals. But if relativists are right, those ideals can never be mistaken, since they just are the ultimate moral standards for each society.

And yet societies are sometimes based on principles of slavery, of war-like aggression, or of sexual, religious, or ethnic oppression. Cultural relativism would turn these founding ideals into iron-clad moral duties, making slavery, sexism, and racism the moral duty of all citizens of those societies. The **iconoclast**—the person deeply opposed to conventional wisdom—would, by definition, always be morally mistaken. This has struck many people as seriously implausible.

But subjectivism is not in the clear here. It faces a similar problem. The cultural relativist makes societies morally infallible—incapable of error—at least with regard to their foundational principles. Yet subjectivists make *each person's basic commitments* morally infallible. True, subjectivism allows that people can make moral mistakes, but only if they fail to realize the implications of their own commitments. When it comes to the basic commitments themselves, subjectivism denies that these can ever be false or immoral.

If morality is in the eye of the beholder, then everyone is seeing things equally well. Millions of people have very sincerely endorsed programs of

ethnic cleansing, male domination, and chattel slavery. Subjectivism turns these prejudices into moral truths.

Moral Equivalence

Subjectivists grant that your moral values, which very likely oppose the ones just mentioned, are also correct. The biased and the bigoted have no monopoly on the truth. Ethical subjectivism is a doctrine of moral equivalence; everyone's basic moral views are as plausible as everyone else's. This can sound liberating and tolerant, and can be put to good use in cutting arrogant people down to size. Such people usually claim to have found the Truth, and often think that they have a special license to force this Truth on others. If subjectivism is correct, the views of such zealots are no better than those of their intended victims.

But they are no worse, either. If ethical subjectivism is correct, then the moral outlooks of Hitler or Stalin are just as plausible as those of a Nobel Peace laureate. And, as we will see in the final chapter (have a look at argument 5, pp. 312–314), if all moral views are on a par with one another, then this is a threat to tolerance, rather than support for it, since those with intolerant outlooks would have a moral view as good as that of their opponents.

Cultural relativists fare a bit better here. They will deny that everyone's moral views are equally plausible. Some people are much wiser in moral matters than others, since some people are better attuned to what their society really stands for. But when it comes to evaluating the basic codes of each society, relativists must allow that every code is equally good. Since the ultimate moral standards are those endorsed by each society, none is better than any other. That may sound egalitarian and open-minded, but what it means in practice is that social codes that treat women or ethnic minorities as property are just as morally attractive as those that don't. That's not an easy thing to accept.

No Intrinsic Value

Here is an ancient moral question: Is something good because we like it, or do we like things because they are good? Ethical subjectivism goes for the first option. There is nothing intrinsically good about promise-keeping, generosity, kindness, or caring. Subjectivists think that these things are valuable, if they are, only because people approve of them. Were our tastes to change, the morality of such actions and character traits would change with them.

That might strike you as suspicious. If it does, then cultural relativism might seem a good alternative. For in that view, moral standards do not depend on the possibly fickle choices of any single person.

Yet cultural relativism faces the same worry. For the relativist, the value of something depends entirely on whether a society's guiding ideals approve of it. When these ideals change, the moral code changes with them. If societies place no value on tolerance, or sexual equality, then in those societies such things have no moral value at all. An open, tolerant society that eventually became a fascist tyranny would not be falling into moral error. If relativism is true, then a society's basic moral ideals (no matter what they stand for) are correct. They are not correct because they measure up to some independent standard. They are correct because a society embraces them.

The problem with such a view is that the ultimate moral principles—whether fixed by each individual or by each society—can be based on prejudice, ignorance, superficial thinking, or brainwashing, *and still be correct*. According to both kinds of relativism, the origins of our basic moral beliefs are irrelevant. No matter how we came by them, the relativist claims that our ultimate moral beliefs cannot be mistaken.

Questioning Our Own Commitments

If subjectivism is correct, then I know what is right so long as I know what I approve of. That's because my approvals (according to subjectivism) are the ultimate test of morality. But what about the situations where I want to know whether my commitments are worthwhile? In these cases, I know what I like, but am still up in the air about its value. This sort of puzzlement *seems* to make sense. I have been in such situations before, where I am unsure of whether I am right to like someone so much, or wrong to be so critical of some action. But if subjectivism is true, this *cannot* make sense, since my approvals and disapprovals are the ultimate test of right and wrong.

The same sort of problem faces cultural relativism. There is no room in this theory to second-guess the guiding ideals of one's own society, since (by definition) they are the correct moral standards of that society. And yet it sometimes does seem to make sense to ask whether the basic principles of one's society are also morally acceptable. If relativism is correct, however, such questioning reveals a confusion about what morality is all about.

Moral Progress

It seems that both individuals and societies can make moral progress. We can do this when our actions become morally better than they used to

be. But I am thinking here of progress in our moral beliefs. This occurs when more of them are true and, in particular, when our most fundamental beliefs change for the better.

The gradual reduction in racist and sexist attitudes in the United States seems to represent this sort of moral progress. The kind of repentant self-examination that German society undertook (and continues to undertake) after World War II also seems a clear improvement over Nazi ideology. When I examine my own life, I see several moral views that I held when I was younger that I now regard as seriously mistaken. I am probably not alone in this.

The problem for relativism and subjectivism is that it does not seem to be able to make sense of the most basic kind of moral progress. If a person's or a society's deepest beliefs are true by definition, then they *cannot* change for the better. They can change, of course. But if subjectivism or relativism is true, then this change cannot represent moral progress.

To measure moral progress, you need a standard. In ethics, that standard is the ultimate moral rule (or rules, if we are pluralists). If subjectivism is correct, that ultimate rule is personal opinion. If relativism is correct, that ultimate rule is given by a society's basic ideals. These cannot be mistaken. If a society gradually eases out of its deeply sexist attitudes, for instance, that cannot be moral progress. That can only be a change to a different moral code. And if relativism is correct, different moral codes are not better or worse than one another. They are morally equivalent.

According to these views, moral improvement is possible only if our more specific moral beliefs line up better with our deepest moral convictions. But these deepest convictions can never improve. No change in them can ever represent moral progress, since they are the ultimate standard by which any such progress could be measured.

If subjectivism is correct, then inmates who experience a deep change of heart while in prison, who adopt new aims of charity and repentance, cannot be showing moral progress. If relativism is correct, then a society that rejects its earlier ideals of racial purity and genocide cannot be making moral progress. That is difficult to believe.

In sum, both forms of relativism encounter some serious difficulties. They make the deepest commitments of each person or society morally infallible, no matter whether such commitments reflect ignorance, bias, sloppy thinking, etc. They are doctrines of moral equivalence, and so deny that compassion, kindness, and benevolence are morally superior to treachery, betrayal, and violence. In rejecting the idea that any

actions or character traits are intrinsically good, they make morality subject to the whims of individuals or societies, as changeable as personal or social opinions. When we search for guidance in examining our most basic commitments, both forms of relativism have nothing to offer. Indeed, they think that such questioning is confused, since (by their lights) things are good only because we value them, and not the other way around. Finally, neither form of relativism can make sense of fundamental moral progress, understood as an improvement in our deepest moral beliefs.

Contradiction and Disagreement

A final problem for both theories is one that you've probably already thought of. It is the problem of contradiction. A contradiction occurs when a statement is said to be both true and false at the same time. It's a contradiction, for instance, to both assert and deny that the Empire State Building is in New York. Theories that generate contradictions are incoherent. They can't be true; they are muddled and inconsistent.

It looks like subjectivism leads to contradiction. For consider its test of truth and falsity:

(S) A moral judgment is true if it accurately reports one's feelings or commitments, and is false otherwise.

If S is correct, then people on opposite sides of a moral debate are both saying something true. The pro-choicer is speaking the truth when saying that abortion is morally right. And the abortion opponent is also speaking the truth when saying that it is immoral. But abortion can't be both right and wrong. That is a contradiction. Since subjectivism generates contradictions, it must be false.

There is a solution to this problem, but it has its costs. The solution implies that we usually don't mean what we say in our moral debates. What we say are things such as:

- The death penalty is immoral.
- Abortion is wrong.
- Eating animals is okay.

But what we *mean* is:

- The death penalty is wrong, according to me.
- I disapprove of abortion.
- As I see it, eating animals is okay.

And just like that, the contradictions disappear! Suppose that you and your friend disagree about whether eating animals is wrong. You say it is; she says it isn't. As the subjectivist sees things, you are saying that you disapprove of meat-eating; she says that she approves of it. These claims don't contradict each other. This sort of strategy will work across the board, for all moral claims, and so we can save subjectivism from contradiction.

Here are the costs. First, subjectivists have to accuse nearly everyone of misunderstanding their own moral claims. And second, such a view eliminates the possibility of moral disagreement.

To illustrate the first problem, consider this conversation:

ME: Genocide is immoral.

SUBJECTIVIST: What I'm hearing is—you disapprove of genocide.

ME: Yes, I disapprove of genocide. But that's not what I'm saying. I'm not talking about my attitudes, I'm talking about genocide. You're changing the subject.

Subjectivists can't make sense of my reply here. It's not that my reply might be false. Rather, my reply is unintelligible, since it assumes that moral talk is about something other than my own commitments. Most of us assume precisely that. If subjectivism is right, we are badly mistaken.

In order to avoid the problem of contradiction, subjectivists have to say that our moral assertions report facts only about our own commitments. When I say that genocide is wrong, I am not saying that it has a certain feature—wrongness. I am saying that I disapprove of it or that my principles forbid it. I am talking about myself. That's not what most people think they are doing when they make their moral judgments.

The second problem is even more serious. Subjectivism is unable to explain the existence of moral disagreement. In order to avoid generating contradictions, subjectivists have to understand all moral judgments as reports of whether I approve of something or not. The claim that meat-eating is wrong becomes the claim that I disapprove of meat-eating. The judgment that bravery is a virtue becomes the claim that bravery is something I admire. And so on. But on this line, moral debates that seem to involve intense disagreement become something completely different. In fact, it now becomes *impossible* for people to morally disagree with one another.

To see this, imagine an earlier dispute.

YOU SAY: It's wrong to eat meat.

AND YOUR FRIEND SAYS: It's okay to eat meat.

The subjectivist translates this as follows:

YOU: I disapprove of eating meat.

YOUR FRIEND: I approve of eating meat.

The contradiction has indeed disappeared. But so has the disagreement. If you are both taking this seriously, you'll agree with your friend's claim, and she with yours. If all that moral judgments do is report people's outlooks, then there is no way to morally disagree with anyone—except to charge them with insincerity. But that seems plainly wrong.

Subjectivism reduces all moral talk to autobiographical reports. Disagreement is impossible, because there is no common subject matter to disagree about. Those who debate the merits of abortion may think that they are discussing abortion, and whether it is has a certain feature—namely, that of being morally acceptable. Not so. They are instead talking about their attitudes of approval and disapproval. If that were true, there would be no point in debating morality at all.

But there certainly seems to be a point (or rather, many points) to engaging in moral discussion. And the appearance of moral disagreement is very vivid. Perhaps such appearances are all illusions. And perhaps moral debate really is pointless. It would indeed be pointless if all we could possibly talk about were our own thoughts about things, rather than about the moral features of the things themselves.

In short, subjectivism faces a dilemma. If we take moral claims at face value, then subjectivism generates contradictions, and so it must be false. If we reinterpret all moral claims to be focused on our attitudes, then the contradictions disappear, but so, too, does moral disagreement.

Cultural relativism faces the same dilemma. It says that *a moral judgment is true just because it correctly describes what a society really stands for*. For instance, if different societies disagree about the appropriate political status of women, then each is speaking the truth when it asserts (or denies) female moral equality. But that is impossible. The statement that women are deserving of full political equality cannot be simultaneously true and false.

Relativists can escape this problem in familiar ways. They will claim that moral judgments are true only relative to social agreements. In this line of thinking, moral judgments are just like legal ones. It isn't contradictory to say that smoking marijuana, for instance, is both legal and illegal, so long as we qualify things to note that it is legal in some areas and illegal in others.

Relativists will say that all of our moral claims have to be understood by reference to social agreements. When you say that meat-eating is right, and your Hindu friend from Calcutta says that it is wrong, what is really being said is:

YOU: Meat-eating is accepted by my social customs.

YOUR FRIEND: Meat-eating is forbidden by my social customs.

And again, both of these claims can be true. The contradiction disappears. There is no single judgment that is both true and false.

But then the existence of cross-cultural moral disagreement also disappears. If the ultimate moral standard is each society's ethical code, and our moral judgments are attempts to describe what our society believes, then the only way to criticize someone is to say that he has mistaken what his society really stands for. It doesn't seem as if that is what serious, engaged moral debate is all about. For instance, it appears possible to note that a society approves of making wives domestic slaves and yet to disagree with the morality of that policy. But that's not so if relativism is to escape the contradiction problem.

So the cultural relativist faces the same dilemma as the subjectivist. Indeed, the relativist is in one way more vulnerable than the subjectivist here. For the cultural relativist may be unable to escape contradiction after all.

People who are members of subcultures—smaller cultural groups located within larger ones—often face a familiar problem. They are forced to choose between allegiance to the larger society and to their particular subculture. They are members of at least two societies, and when their ethical codes conflict, these unfortunate people are faced with contradictory moral advice.

This isn't some philosopher's fiction. Such cases happen all the time. We could easily multiply examples, but this famous one from my home state should be enough to make the point.

Consider the facts of *Wisconsin v. Yoder*, a case resolved by the U.S. Supreme Court in 1972. Wisconsin then required regular school attendance of all children up to the age of sixteen. The sons of three Old Order Amish families had stopped going to school after the eighth grade, in obedience to their parents' beliefs that continued schooling would conflict with their religious values. The students were found guilty of violating the state law, but the verdict was overturned by the State's Supreme Court.

Wisconsin then appealed to the U.S. Supreme Court, which sided with the Amish families.

In its decision, the Court's majority announced that:

They [the Amish families] object to the high school, and higher education generally, because the values they teach are in marked variance with Amish values and the Amish way of life; they view secondary school education as an impermissible exposure of their children to a "worldly" influence in conflict with their beliefs. The high school tends to emphasize intellectual and scientific accomplishments, self-distinction, competitiveness, worldly success, and social life with other students. Amish society emphasizes informal learning-through-doing; a life of "goodness," rather than a life of intellect; wisdom, rather than technical knowledge; community welfare, rather than competition; and separation from, rather than integration with, contemporary worldly society.

The schoolchildren lived in (at least) two societies at once: their Amish community and the larger state of Wisconsin. If relativism is correct, then the morality of your actions depends entirely on whether they are allowed by the standards of the society they are performed in. But if you live in different societies, and their ethical codes clash, then your actions will be both moral and immoral. That is a contradiction.

We could solve this problem if we could figure out which society's code is more important. But relativism doesn't allow us to do that. By its lights, no society's moral code is any better than another's. We might be tempted to let the children decide, and say that the social code that takes priority is the one that the children prefer. But this would undermine cultural relativism, since such a move would make the morality of their actions depend on personal choice. They would get to pick the code that is to govern their lives. That is subjectivism, not relativism.

Indeed, critics of cultural relativism often say that the doctrine eventually collapses into subjectivism. When your views and society's views clash, why think that society is always right? If morality is created by humans, then it is hard to justify the claim that moral wisdom always lies with the masses rather than with individuals. The majority may have the power to force the minority to do as it says. But might doesn't make right.

Subjectivists claim that in conflicts between personal and social commitments, the individual is always morally wiser. Cultural relativists take the opposite line. But perhaps things are not so cut and dried. Sometimes individuals have the upper hand; sometimes societies do. And sometimes,

perhaps, both individuals and societies are mistaken, even in their deepest commitments. If that is ever so, then we must look elsewhere for an account of morality's true nature.

Ideal Observers

There is a natural way to fix some of these problems for the subjectivist and relativist. We should guarantee that those who create the moral law (whether each individual or whole societies) are not choosing from ignorance, but are equipped with full information. We should also make sure that they are reasoning clearly and avoiding logical errors. In other words, rather than allow us as we actually are (warts and all) to have the final word in morality, we should make the desires and choices of **ideal observers** the ultimate standard of morality. Ideal observers can survey the scene more dispassionately, more knowledgeably, more rationally. They are better suited to inventing the moral law than we mere mortals are.

According to this new and improved version of subjectivism, an act is morally right just because I *would* favor it were I fully informed and perfectly rational. The relativist version says that acts are morally right just because a society *would* approve of them were its members fully informed and rational.

This will surely correct some of the problems that we have noted: (1) Even the core moral beliefs of individuals and societies may now be mistaken, as their views may fail to measure up to those of the ideal observers. (2) Further, the views of individuals and societies will *not* be morally equivalent, since some will more closely approximate those of the ideal observers. (3) The sincere endorsements of slavery and genocide will not automatically be morally authoritative, since such endorsements are almost always based on ignorance and irrationality. (4) Moral progress will now be possible, and will occur when the moral views of individuals and societies more closely reflect the attitudes of ideal observers. (5) There will be real disagreement between conflicting moral views, since moral judgments will not be reports of personal opinion or cultural consensus, but will rather be claims about what ideal observers will approve of.

These are real improvements. But ideal observer views are not problem-free. In fact, there are two serious concerns. The first occurs if there is ever any disagreement among ideal observers. The ideal observer view says that perfectly rational and intelligent people create morality through their choices. If that is so, then if such people make conflicting

choices, this will cause contradictions. And contradictions fatally undermine any theory that contains them.

Perhaps perfectly smart and rational people will never disagree about anything. But why the optimism? Those who know all there is to know about embryology, for instance, might still morally disagree about abortion. After all, in the ideal observer views, such geniuses are not trying to understand the morality of the actions they are assessing. Before they make their decisions, there is no morality. Ideal observers don't respond to a world with moral features. Their preferences and choices create morality. But then there doesn't seem to be anything to prevent them from having conflicting attitudes. If they do, contradiction results.

I think that there is a successful solution to this problem. We can borrow a strategy we've seen before,³ when discussing a similar problem that arose for social contract theories and virtue ethics. The strategy tells us that an action is morally required or forbidden only if *all* ideal observers agree in their attitudes about it. If all ideal observers endorsed an action, then it would be morally required. If they all opposed it, it would be forbidden. And if they disagreed on the matter, then it would be morally permitted—neither required nor forbidden. By making morality depend on the attitudes of all ideal judges, rather than each one individually, this theory can indeed avoid contradiction.

But another problem cannot be handled so easily. The view on the table says this: Nothing is intrinsically right; things become right just because an ideal observer would favor it. But what if such people thought that killing off the mentally ill was a great idea? What if they thought that sadism was preferable to compassion? What if they approved of apartheid policies? You might think such a thing impossible. But why? Evil people need not be factually ignorant or illogical. Vast knowledge doesn't guarantee a sympathetic nature. Greater logical skills don't automatically translate to greater kindness. Even the most rational and well informed among us can be biased, hateful, and cruel.

Recall how we got here. Subjectivism and cultural relativism allow the basic views of individuals or societies to determine the ultimate moral standards. But such basic views can be the product of ignorance, bias, and poor reasoning. We tried to fix this problem by changing the theories so that the authors of morality were ideal versions of us. They would be people with perfect information, and perfect logical skills. And yet, as we've seen, this modified view has troubles of its own, and fails to solve the worry that led to its creation. The very smartest people can also be the coldest and cruelest.

This is a deep problem for ideal observer views. In fact, the problem should be a familiar one, since it is the same one that threatens the divine command theory.⁴ The divine command theory says that no acts are intrinsically right or wrong; their morality depends entirely on whether God approves of them. In this view, acts are morally right just because God insists that we do them, and wrong because He loathes them.

The basic problem is that actions don't become right just because someone (even God) happens to favor them. Think back to our earlier discussion of the Euthyphro dilemma.⁵ Either God has reasons for His commands or He doesn't. If He doesn't, then the commands are arbitrary, and can't provide the basis for a legitimate morality. But if God does have reasons for His commands, then these reasons, rather than God's say-so, are what explain why various actions are right. God can ratify the moral standards. He can know every one of them. He can convey them to us. But He cannot be their author, on pain of resting morality on arbitrary foundations.

The same line of reasoning works to undermine all of the views we have considered in this chapter. Subjectivism, cultural relativism, and ideal observer theories all share the same basic structure. On these views, nothing is right or wrong in and of itself. Actions have the moral status they do only because I or my society actually approve of them, or would approve of them if we were perfectly intelligent. How can the decisions of any such person or group be so powerful as to transform a valueless activity into something good or right?

These morally all-powerful people either are or aren't basing their decisions on good reasons. If there are no good reasons to back up their decisions, then the decisions are arbitrary, and cannot be the basis of a morality worthy of our respect. But if there are good reasons to back up the decisions, then the reasons, rather than the decision, determine the morality of the actions in question.

Suppose, for instance, that I (or my society or an ideal observer or God) have reasons that support my disapproval of torture. And these are the reasons: the pain it imposes, its unreliability as a source of valuable information, the disrespect it reveals, and the way it renders its victims utterly powerless. If these really are good reasons, then they are all that's needed to make torture wrong. My disapproval doesn't add anything to these reasons. If I am really wise, then my disapproval can be very good *evidence* of something's immorality. But the approval cannot turn a morally neutral action into a forbidden one.

Socrates's argument against the divine command theory is just as powerful when brought against subjectivism, relativism, and ideal observer theories. If his line of reasoning is correct, then personal approval is not enough to make something right. Acts are right because they are supported by excellent reasons, and not because individuals or groups just happen to favor them.

Conclusion

Both cultural relativism and ethical subjectivism are popular ways of challenging the idea that morality is objective. But as we've seen, both theories face a similar set of problems. They make all moral views or all social codes morally equivalent. They make the deepest commitments of each person or each society morally infallible, even if the commitments are based on ignorance or prejudice. Neither theory offers a way to evaluate our guiding ideals, since these ideals are correct because we endorse them, and not the other way around. Neither theory allows for fundamental moral progress. Both theories generate contradictions, and can eliminate this worry only by making moral disagreement impossible.

This laundry list of complaints explains why cultural relativism and ethical subjectivism have found little favor among philosophers. For those with doubts about the objectivity of morality, nihilistic alternatives may have more to offer.

Notes

1. The term "moral skepticism" sometimes refers to the view that gaining moral knowledge is impossible. (That's the way I used the term in chapter 16, for instance.) I am going to use the term here in a different way, noted above—namely, to refer to all theories that deny the existence of objective moral standards.
2. All terms and phrases that appear in **boldface** are defined in the glossary at the end of the book.
3. See chapter 14; see also chapter 17, pp. 254–256.
4. See chapter 5 for a more detailed discussion of the divine command theory.
5. See pp. 60–65, and the related discussion in chapter 17, pp. 256–258.

Moral Nihilism

There are basically three options when it comes to determining the status of morality. Morality might be objective, its rules and its demands applying to us independently of our opinions and desires. Or it might be relativistic, and depend for its authority on personal or cultural tastes and choices. Finally, morality may simply be a kind of make-believe, a complex set of rules and recommendations that represents nothing real. This last option is known as **moral nihilism** (from the Latin word *nihil*, meaning “nothing”).

Moral nihilists join with relativists in opposing ethical objectivism. Morality is wholly a human creation—in this, nihilists and relativists are united. But nihilists are no fans of ethical relativism. Relativists believe in moral goodness, moral duty, moral virtue. Nihilists don’t. Nihilists deny that there are any moral qualities. There are no moral requirements. Nothing is morally good. Nothing merits praise or blame.

There are two important forms of moral nihilism: the **error theory** and **expressivism**. Error theorists claim that our moral judgments are always mistaken. Expressivists deny this, while also denying that our moral claims can ever offer an accurate take on reality. Expressivism is the more complicated doctrine, so let’s ease our way into the nihilist camp by first considering the error theory.

Error Theory

Did you ever have the feeling, deep down, that morality is a sham? That it’s just a set of traditional rules inherited from ancestors who based it on

ignorance, superstition, and fear? Perhaps it's only a convenient fiction, with no underlying authority at all.

The error theory of morality is built upon these doubts. It is defined by three essential claims:

1. *There are no moral features in this world.* Nothing is morally good or bad, right or wrong, virtuous or vicious, etc. A careful inventory of the world's features will reveal all sorts of scientific qualities: being symmetrical, being a liquid, being two feet long, carbon-based, spherical, etc. But the list will contain no moral features.

2. *No moral judgments are true.* Why not? Simple: there is nothing for them to be true of. There are no moral facts. And so no moral claims can be accurate, since there are no moral facts for them to record.

3. *Our sincere moral judgments try, and always fail, to describe the moral features of things.* Thus we always lapse into error when thinking in moral terms. We are trying to state the truth when we make moral judgments. But since there is no moral truth, all of our moral claims are mistaken. Hence the error.

It follows that:

4. *There is no moral knowledge.* Knowledge requires truth. If there is no moral truth, there can be no moral knowledge.

Error theorists are not launching some small-scale attack on morality. They are not criticizing our current views on, say, welfare policy or capital punishment, and trying to replace them with better ones. Rather, as they see it, *all* moral views are equally bankrupt. There is some very deep mistake that everyone committed to morality is making. The error theorist promises to reveal that mistake, and so to expose the real truth: Morality is nothing but a fiction.

For those who are fond of analogies, the following may help. The error theory is to morality as atheism is to religion. Error theorists and atheists are skeptics. They deny the truth of a widely accepted world view. They do this by trying to pinpoint an error that is said to lie at the very heart of the system they oppose.

Atheists are, in effect, error theorists about religion. They believe that there are no religious features of the world, that no religious claims are true, and that religious believers try (and always fail) to speak the truth about God. They deny that there is any religious knowledge. The central explanation for all of this is simple. If atheists are right, then common religious claims (God speaks to me; God created the universe; God knows everything) are

all wrong, because they are all based on a mistaken assumption. The basic error of religion is its assumption that God exists. He doesn't.

Of course, the atheist will muster arguments for that claim, but we cannot pause here to consider them. Atheists must also provide an account of why so many hundreds of millions of people have come to believe religious claims. Marx spoke of religion as the opiate of the masses—an invention designed by the powerful to keep the powerless in check. Freud regarded religion as an illusion designed to ease our insecurities and fears. There are many such tales to tell, but we should leave the telling to psychologists, anthropologists, and political scientists. It is the philosopher's job to determine whether the basic error lives up to its name.

Error theories about morality must also explain why so many people so strongly believe in moral virtue, goodness, and duty. As before, such explanations are fairly easy to come by. Morality helps us to survive; morality gives us a sense of purpose and meaning; we are hard-wired to believe in it, etc. Again, we must leave the assessment of such accounts to others. Our goal is to determine whether there is even any need for them, by asking whether morality really is based on an error.

And that depends on what the fundamental error of morality is supposed to be. In principle, we can develop any number of error theories, depending on which basic error morality is supposed to commit. But in practice, there really has been only one candidate.

All error theorists have agreed that the fatal mistake that undermines morality is its assumption that there are objective moral standards that supply each of us with an excellent reason for obedience, regardless of what we care about. According to error theorists, just as religion crucially depends on the supernatural, morality essentially depends on its being objective and providing us with **categorical reasons**—reasons that apply to us regardless of whether acting on them will get us what we want.¹ If this central assumption is mistaken, then the entire enterprise of morality is bankrupt.

There are two substantial points that error theorists must convince us of. First, they must show that buying into morality really does assume a commitment to moral objectivity and categorical reasons. That will be news to many—to subjectivists and relativists, for instance, and to expressivists, whose views we are soon to discuss. If morality does not, in fact, rely on these assumptions, then the error theorist's criticisms will fail.

But suppose that the coherence of our moral thinking and practice does indeed depend on the twin assumptions that morality is objective, and provides us with categorical reasons. This reveals the second burden

that error theorists must shoulder: they must show that at least one of these assumptions is false.

Perhaps they can do that. We have already considered arguments concerning our reasons to be moral,² and will (in the next chapter) look at the most prominent attempts to undermine the objectivity of morality. So, rather than repeat those efforts, let us consider some implications of the error theory.

Though very few people outside of philosophical circles have ever heard of the error theory, the worry it expresses is familiar enough. And so, too, is the typical response: Abandoning morality would have absolutely terrible results. Further, to seriously consider that morality may be a fiction is to show a corrupt mind, and error theorists are therefore not to be trusted. Once people give up on morality, they will feel free to act in any way they please. It won't be pretty.

There are actually two lines of criticism at work here, and both are mistaken. The first we might call the *Argument from Disastrous Results*:

1. If widespread acceptance of a view would lead to disastrous results, then that view is false.
2. Widespread acceptance of the error theory would lead to disastrous results.
3. Therefore, the error theory is false.

It's an interesting question whether premise 2 is true. I won't venture an opinion here on whether it is. Nor do we need to, since premise 1 is definitely false. The truth of a theory does not depend on the results of everyone's embracing it.

To see this, consider a similar argument against atheism. Some people claim that widespread belief in atheism would generate disaster. I don't know if that's right, but suppose it is. Still, that is no evidence that God exists. After all, if some highly classified state secrets were widely publicized, that might cause disaster, too. But that hardly shows that the claims made in those documents are false. So we cannot undermine the error theory by arguing that its popular acceptance would lead to the downfall of civilization—even if it would.

Some have found premise 1 tempting because they have confused it with a close cousin:

- 1a. If widespread practice of a given action will lead to disastrous results, then that action is immoral.

But this really is a confusion, since 1 speaks of theories and their truth, rather than actions and their moral status. And in any event, 1a is false, as we saw in an earlier discussion.³ If everyone were to practice celibacy, then disaster would ensue. That doesn't show that it is immoral to be celibate.

Others have confused 1 with a more specific version:

- 1b. If widespread acceptance of a *moral theory* would yield disastrous results, then that theory is false.

This cannot be so easily dismissed. Many philosophers have thought that 1b is true, though a number of others—mostly utilitarians—have rejected it.⁴ Fortunately, we do not have to settle the matter here, because *the error theory is not a moral theory*. It does not try to tell us where our duty lies, or which character traits are virtues. The error theory rejects *all* moral theories, and says that every single one is mistaken. Since the error theory is not a moral theory, principle 1b simply does not apply to it.

If the error theory isn't a moral theory, then what is it? In fancy terms, it is a **metaphysical** theory—a theory about what the world is truly like, and what really exists. Theism is a metaphysical theory. It says that God exists. Other metaphysical theories try to defend the existence of the soul, or free will, or immortality.

The basic problem with the Argument from Disastrous Results is that metaphysical theories cannot be tested in the way that its first premise claims. For instance, we cannot prove that we have free will, just by showing (if we can) that terrible results would occur if we abandoned our belief in it. Metaphysical theories try to tell us what the world is like. Such theories might contain some bitter truths, ones that, if widely accepted, would lead to heartbreak, or loss of faith, or the breakdown of longstanding customs and social practices. (That's what makes them bitter.) At best, this might give us some reason not to publicize these claims. But that is no reason to suppose that they are false.

Another popular objection to the error theory is really targeted at error theorists themselves. The idea here is that those who reject categorical reasons and the objectivity of morality are bound to be untrustworthy. Conscience serves as an effective check on our antisocial impulses only if we see moral duty as something real, as a set of rules imposed from the outside, ones that have genuine authority over us. Error theorists reject this picture, and so will feel free to let their destructive, self-interested impulses take control of their decisions. Since error theorists are so untrustworthy, the views they put forth are not to be trusted, either.

But this way of thinking is mistaken as well. Error theorists can care deeply about others, and can be strongly opposed to doing the things that we traditionally regard as morally wrong (killing, raping, stealing). Of course error theorists, if consistent, will not regard such actions as immoral. But they may still be dead set against such behavior; they might find it offensive, undesirable, unproductive, or otherwise unappealing.

Still, one might argue that error theorists are not *guaranteed* to have the sorts of goals that we associate with upright behavior. And that is true. But then again, such a guarantee fails to hold of many people who reject the error theory. How many atrocities each year are committed in the name of one morality upon those with different moral beliefs? Far more than are committed by error theorists!

All that aside, this kind of criticism does nothing to address the issue of whether the error theory itself is true. Indeed, it represents a classic kind of fallacy—the *ad hominem* attack. When leveling such a critique, one tries to undermine the truth of a position by criticizing the character of its supporters. Don't like the message? Attack the messenger. Aside from being bad sport, such a strategy entirely misses the mark. We want to know whether morality is all make-believe. We can't answer that question by engaging in character assassination.

The only way to answer it is by doing two things. First, we need to determine whether error theorists are correct in thinking that morality really does depend on two assumptions: (1) that it is objective, and (2) that it supplies reasons for everyone to obey it, regardless of personal desires. I am not sure about whether this is what we really are committed to when thinking morally. Certainly, subjectivists and relativists deny these assumptions. And so they will deny the existence of any error, since they reject the thought that morality depends on those assumptions.

But suppose that error theorists are right about what we are committed to when we adopt a moral outlook. To defend their view, they must then show that categorical reasons do not exist, and that morality is not objective. Can they do this? Stay tuned.

Expressivism

Expressivism is another family of views that deny the objectivity of morality. Indeed, expressivists accept the first two claims that define the error theorist's point of view:

1. There are no moral features in this world.
2. No moral judgments are true.

Unlike error theorists, however, expressivists think that morality is in pretty good shape. There is no deep error at the heart of our moral thinking. But how can that be, if 1 and 2 are true?

Simple. Expressivists reject the third feature that is essential to error theory:

3. Our sincere moral judgments try (and always fail) to describe the moral features of things.

According to expressivists, we are not trying to speak the truth when making moral judgments. We are not making an effort to describe the way the world is. We are not trying to report on the moral features possessed by various actions, motives, or policies. Instead, we are venting our emotions, commanding others to act in certain ways, or revealing a plan of action. When we condemn torture, for instance, we are expressing our opposition to it, indicating our disgust at it, publicizing our reluctance to perform it, and strongly encouraging others not to go in for it. We can do all of these things without trying to say anything that is true.

One of the basic ideas behind expressivism is that moral claims function very differently from straightforward factual claims. Factual claims try to represent the way the world really is. If expressivism is right, moral claims serve quite different purposes.

I know that sounds puzzling. To get a better sense of what is going on here, consider this sentence:

(A) Torture is immoral.

It appears to function just like the sentence

(B) The sea is salty.

Sentence B tells us that the sea has a certain feature—being salty. Sentence A looks similar. It tells us that torture has a certain feature—being immoral. And there's nothing special about A. All moral claims seem to assign a moral quality to something or other. But if there aren't any such qualities, then aren't our moral claims always mistaken?

Not necessarily. If expressivism is true, then the similarity between sentences such as A and B is only superficial. When we say that torture is immoral, for instance, we are not *describing* torture. We are *not* saying that

it has any features at all. We aren't even describing our feelings about it (as subjectivists claim). Rather, it's as if we were saying:

- Torture—argghhh! *or*
- Don't torture! *or*
- Let me plan a life that doesn't include torturing others, *or*
- Won't everyone please refrain from torture?

These utterances can't be true. But they can't be false, either. And that's the central difference between expressivism and error theories. The error theorist thinks that our sincere moral claims are always meant to state the truth, but since there isn't any moral truth, such claims are all mistaken. The expressivist, by contrast, thinks that our moral claims are largely alright, since they are doing what they are intended to do. And what is that? Moral claims are not in the business of holding up a mirror to the world. Their job is to vent our feelings, give orders and commands, and express our commitments. Since they manage to do that just fine, there is no reason to charge them with error.

The basic philosophical motivation behind expressivism is pretty straightforward. Expressivists want a way to have confidence in morality while rejecting ethical objectivity. In doing so, they also want to avoid the difficulties that hamper cultural relativism and ethical subjectivism.

The biggest problem for relativism and subjectivism (as we saw in the previous chapter) is that these views either generate contradictions or are unable to explain moral disagreement. Expressivists handle both problems with ease.

Contradictions arise when the same claim is said to be true and false at the same time. If expressivists are right, no moral claim is either true *or* false. And so moral contradictions disappear.

Expressivists see moral disagreement as a clash of emotions or personal commitments. Debates about torture, for instance, reveal nothing about torture's moral features (since there aren't any), but a lot about the feelings of the differing parties. One side feels angered and upset by torture, and the other doesn't. The emotional investment we see in so many moral debates is just what we should expect, given the expressivist analysis of moral disagreement.

Expressivists cannot escape the worry about moral equivalence, however. Their official view is that there are no moral values, and no moral truth. If that is so, then all moral views are on par with one another. Some such views may be more internally consistent than others; some may

contain fewer factual errors; others may be more likely to bring happiness or contentment. Yet none of these things makes an outlook morally better than another. None makes a moral view closer to the truth.

There are further worries. Three seem to be especially serious.

How Is It Possible to Argue Logically About Morality?

If, as expressivists say, moral claims cannot be true, then this makes it very difficult to understand how moral argument is possible. Logical argumentation is *truth-preserving*—a logically valid argument is defined as one whose conclusion must be true, provided that its premises are true.⁵ If, as expressivists say, moral claims cannot be true, then how could they possibly be used to support other claims?

To see the worry here, consider this argument:

1. All actions that dehumanize people are immoral.
2. Torture dehumanizes people.
3. Therefore, torture is immoral.

The argument appears to be logically perfect. If you accepted both premises, you would have to accept the conclusion. After all, this argument has the very same logical structure as a philosophical classic:

1. All men are mortal.
2. Socrates is a man.
3. Therefore, Socrates is mortal.

And everyone admits that *this* argument is logically valid. Given the truth of the two premises, the conclusion can't possibly be false.

From a logical point of view, these two arguments appear to be identical. And there is a natural explanation for this: The first two premises in each argument can be true, and if they are, then the conclusion must be true as well. But expressivists cannot accept this natural explanation, since they deny that moral claims can be true.

Look at the first argument. If premise 1 really means:

- 1a. Dehumanizing actions—yecchhh! *or*
- 1b. Don't dehumanize people!

then there is no way that this premise can be used to logically support any conclusion. But it seems clear that it is being used in precisely this way. The logic of the first argument is watertight—even if you are suspicious about the merits of the actual premises. Indeed, logical moral

argumentation seems like a real possibility—we do it all the time (and have been doing it throughout this book). Expressivism does not seem able to explain this.

Expressivism and Amoralists

An amoralist is someone who sincerely makes moral claims, but is entirely unmoved by them. Such people create a serious problem for expressivism.⁶

Expressivists warn us not to be fooled by the superficial similarity between factual claims (the sea is salty) and moral ones (torture is immoral). Moral claims assert nothing. They describe nothing. Instead, they express our feelings. Indeed, that is how the expressivist explains why our moral judgments so reliably motivate us. These judgments convey our feelings, and our feelings are what move us to act.

But this makes it impossible for someone, say, to really think that charity is admirable and yet be indifferent when it comes to forking out his own money. It would be impossible for a soldier to think it his duty to face enemy fire, while remaining completely unmotivated to do so. Such cases may really be impossible. But the evidence points the other way. Amoralists are unusual, to be sure, but not unheard of.

Indeed, expressivism faces the same problem that psychological egoism does.⁷ In the face of evidence that supports the existence of altruism, egoists insist that people must be either deceiving themselves or lying to us about their motivations. Expressivists have to say the same thing about the evidence of amoralism. All such evidence must be disqualified. The person who really is lacking in motivation cannot be sincere in his moral claims. Or, if he is sincere, then he really must be motivated, and so his claims to the contrary are either lies or instances of self-deception. Such a diagnosis may be correct. But the burden here weighs heavily on the expressivist's shoulders.

The Nature of Moral Judgment

A final concern has to do with whether expressivists are right to challenge the dominant view of what we do when we make moral judgments. Ethical objectivists, ethical relativists, and error theorists agree on almost nothing. But they have reached consensus on one point: Moral claims try to report the moral features that actions (or people or policies, etc.) actually have. Moral judgments can be true or false, depending on how well they represent the truth about which things have which moral qualities.

Expressivists deny this. They reject the idea that moral claims are trying to represent the way things are. They deny that there are any moral features. They deny that moral judgments could ever offer accurate reports about reality. How do we know whether their hypothesis is correct?

Return to our original example, the claim that torture is immoral. If we read this literally, the sentence says that torture has a certain feature—being immoral. As we saw, expressivists can't read it this way. They have to paraphrase this sentence so that it isn't assigning any specific feature to torture. Perhaps you found the various translations they offered (p. 299) to be attractive. But what about these claims:

- Nobody but Jeff knows how to behave when the boss is around.
- I'm not sure whether torture is ever acceptable, but I am sure that some who are wiser than I am have the correct answer.
- There is a difference between an action's being required, morally good, virtuous, and deserving of praise.
- Some actions fulfill moral duties and yet lack moral worth.
- The degree of punishment should match the degree of wrongdoing.
- Virtue is its own reward.
- If war is immoral, then military generals are less virtuous than they seem.

Read as objectivists, subjectivists, relativists, or error theorists would do, the sentence structure of these claims is transparent. They are readily understandable. It's not at all clear how expressivists could reword them to turn them into commands, emotional expressions, or plans.

People put their words to various purposes. The best way to tell whether people are joking, insinuating, inviting, or trying to state the truth is simply to ask them. People are usually pretty trustworthy on this score. (Not always, of course. When we play to win at poker, or negotiate a business deal, we may try to deliberately mislead the ones we're with. And jealousy, shame, fear, or anger can sometimes blind us to our true motives.) Still, exceptions aside, people are the best judges of their own intentions, and their testimony in such matters is usually reliable.

That's not a thrilling point, but it has a direct bearing on expressivism's plausibility. For when we ask people how they think of their moral claims, almost everyone will reject the expressivist analysis. For the most part, we do regard our moral claims as true. We regard our opponents' views as false. We take our moral condemnations to be cases of describing, say, the injustice of insider trading, or reporting the moral corruption of a vicious tyrant.

At bottom, we intend our moral judgments to function as something other than emotional outbursts or expressions of commands or plans.

We might all be lying, or deceiving ourselves about what we are actually doing. But by far the more charitable view is that we mean what we say. When making moral judgments, we are trying to speak the truth. We are intending to state the facts. We aim to accurately present the moral details of the situations we are thinking of. If that is so, then expressivism is in serious trouble.

Conclusion

The vision of morality that so many of us believe in—one that sees morality as a set of objective duties and rules, supplying each of us with strong reasons to do as it says—may be fundamentally mistaken. Subjectivists and relativists certainly think so. And so do error theorists and expressivists. Error theorists are the most cynical of the lot, thinking as they do that morality is a bankrupt enterprise. This thought stems from their view that our moral thinking is based on assumptions that turn out to be false. If the foundations of an entire way of thinking are corrupt, then the whole world view must come tumbling down. That's the way atheists see religious claims. And that is the way error theorists see moral ones.

Expressivists are not so pessimistic. They agree with error theorists in denying that ethics is objective, and denying that moral duties supply us with categorical reasons. But since expressivists do not believe that moral thinking rests on these foundations, they don't feel the need to accuse the rest of us of error. Moral talk does pretty much what it is supposed to do—give vent to our emotions, express our feelings about things, and signal our commitments. And so there is little reason to worry.

But as we have seen, expressivism does give us some cause for concern. True, it offers a picture of morality that frees it of fundamental error. It solves the problem of contradiction. It neatly explains the nature of moral disagreement. It supports our view that emotions are a central part of moral judgment. And yet it has difficulty making sense of how logical moral argumentation is possible. It fails to make room for amorality. It clashes with our views about what we are intending to do when we make moral judgments.

Expressivists are hard at work on these problems. There is a very lively conversation about the pros and cons of expressivism going on right now in philosophy departments around the world. In the last decade, this conversation has reached levels of sophistication that would have been

unthinkable just a generation ago. It's far too early to tell whether expressivism is down for the count, or whether its defenders can identify new solutions to these perennial problems.

Even if, in the end, expressivism turns out to have more costs than benefits, this doesn't leave objectivists in the clear. It may be that morality is all make-believe, as error theorists claim. Most of us (especially we textbook authors!) hope this isn't so. Yet a hope is hardly evidence one way or the other. Whether we should think of morality as bankrupt depends on what the best arguments tell us about the nature of morality. We'll be in a better position to see the force of those arguments after working through the next chapter.

Notes

1. For more on categorical reasons, please see the discussion in chapter 11, pp. 152–55.
2. See chapter 8, pp. 102–105, chapter 11, pp. 152–155, and chapter 14, pp. 191–195.
3. See chapter 11, pp. 146–147.
4. Many utilitarians think that widespread acceptance of the principle of utility might actually have quite bad results, despite firmly believing that the principle is true. See chapter 10, pp. 131–132 for a discussion of this point.
5. For more on logic and validity, please see the Introduction, pp. 7–8.
6. For more on the amoralist, see the discussion of Hobbes's Fool in chapter 14, pp. 191–195. See also chapter 21, pp. 318–319.
7. See the discussion in chapter 7, pp. 98–99.

Ten Arguments Against Moral Objectivity

Ethical objectivism is the view that there are some objective moral standards. Given my understanding of objectivity, this amounts to the view that these standards apply to everyone, even if people don't believe that they do, even if people are indifferent to them, and even if obeying them fails to satisfy a person's desires. Moral claims are objectively true whenever they accurately tell us what these moral standards are, or tell us about what these standards require or allow us to do.

Moral skepticism, as I have defined it here, is the view that ethical objectivism is false, and thus that there are no objective moral rules and no objective moral truths. Unsurprisingly, the pros and cons of ethical objectivism and moral skepticism are mirror images of one another. As with any two contradictory positions, a big reason to favor one side is unhappiness with the other. Perhaps the biggest reason that so many people are moral skeptics is a suspicion that ethics just cannot be objective.

Naturally, objectivists are happy to return the favor, and usually defend their own position by raising their serious doubts about moral skepticism. We've considered the leading skeptical views in the last two chapters, and have remarked on some of the difficulties they face. If these worries cannot be solved, and if objectivism can be defended against criticisms, then objectivism wins by default. Relativism, nihilism, and objectivism are the three options when it comes to the status of ethics. If two of these can be defeated, then the one left standing must be the correct account.

Objectivists, of course, believe that they'll be the ones standing at the end of the day. Whether they are right about that depends on their ability

to handle the many criticisms that have been sent their way. It's now time to focus on these criticisms, to see whether objectivism can respond to them in a satisfying way.

There are many sources of doubt. As we'll see, some of the most popular arguments are also the least plausible. But others represent deep and serious challenges. With the hardest of the objections, it is impossible to offer a final verdict in just a page or two, which is all I will allow myself here. My goal in this chapter is simply to show that, despite widespread doubts about ethical objectivism, none of the most popular skeptical arguments is obviously correct, and some, indeed, are pretty plainly unacceptable. And to those that represent more significant challenges, there are potentially promising replies that objectivists can offer.

Let's consider some of the least plausible arguments first, before turning to critiques of objectivism that are more difficult to handle.

1. Objectivity Requires Absolutism

Many people claim that if morality were objective, then moral rules would have to be absolute. And since they aren't, morality isn't objective after all.

The *Argument from Absolutism* summarizes this line of thought:

1. If moral claims are objectively true, then moral rules are absolute.
2. No moral rule is absolute.
3. Therefore, moral claims are not objectively true.

An absolute moral rule is one that is always wrong to break—no exceptions.

I don't know if there are any absolute moral rules.¹ If there are, good candidates would include the prohibition on rape and on deliberately killing innocent people. Luckily we don't have to settle this issue here, because even if premise 2 is true, and there are no absolute moral rules, premise 1 is false.

That first premise tells us that if moral standards are objective, then every moral rule is absolute. But that isn't so. The moral rule that forbids us from lying is probably not absolute; in some cases, morality would probably allow us to lie. For all we know, though, that rule could be objective. Ross thought that the fundamental moral rules are objective.² But he denied that they are absolute. If God exists, and creates or reveals the moral law, then morality would be objective. But God might allow us to lie in certain circumstances, and might also permit us (in unusual cases) to

break other moral rules. There is nothing in the very idea of an objective morality that requires moral rules to be absolute.

There is a general reason for this. The objectivity of moral rules has to do with their *status*: with whether they are ever true and, if so, with the role of human beliefs and desires in fixing their truth. The absoluteness of moral rules has to do with their *stringency*: with whether it is ever okay to break them. There is no direct connection between matters of status and stringency. This is clear when it comes to natural laws. Various biological and psychological laws admit of exceptions, and so are not absolute, even though they are objective.

This does not of course show that moral rules are objective. But it does support the view that even if they are, they do not have to be absolute. So premise 1 is false. And since it is, this argument does not threaten ethical objectivism.

2. All Truth Is Subjective

A popular thought in some circles is that claims can be true only relative to individual perspectives. On this line, there are no objective truths at all. Forget about morality for a moment: Claims in logic, chemistry, or history can never be objectively true, either. So it's no surprise that objective morality is an illusion. The *Argument Against Objective Truths* couldn't be simpler:

1. There are no objective truths.
2. Therefore, there are no objective moral truths.

The first thing to note about this argument is that, if it works, then there is no special problem for morality. Most moral skeptics are trying to show that morality is in some ways second-rate, that it fares poorly in contrast to more rigorous disciplines such as mathematics and physics. This argument abandons that strategy in its embrace of a global kind of skepticism.

The problem with this argument is its premise.³ Premise 1 is either true or false. If it is false, then the argument crumbles right away. So suppose that it is true. But this is impossible. The premise *cannot* be true. For if it were, then there would be at least one objective truth—premise 1. And if there is at least one objective truth, then premise 1 is false! No matter how we look at it, then, this premise is false.

Since that is so, it follows directly that there are at least some objective truths. Perhaps none of them is a moral one. But we can't rely on this argument to support that skepticism.

3. Equal Rights Imply Equal Plausibility

I have heard countless moral disputes end on this conciliatory note: "Well, everyone has a right to their opinion. You have your view, and I have mine. Maybe we're both right."

This familiar refrain is sometimes taken one step further in the following way: Since everyone has a right to a moral opinion, no one's moral views are any better than anyone else's. And if everyone's moral opinions are on par with one another, then there is no objective moral truth. These thoughts can be combined into an *Argument from Equal Rights*:

1. If everyone has an equal right to an opinion, then all opinions are equally plausible.
2. Everyone has an equal right to his or her moral opinions.
3. Therefore, all moral opinions are equally plausible.
4. If all moral opinions are equally plausible, then ethical objectivism is false.
5. Therefore, ethical objectivism is false.

The fourth premise is true. No question about it. If moral standards are objectively correct, then some people's views are going to be very far from the mark, and others are going to be right on target.

I also believe that the second premise is true. Everyone has a moral right to freedom of conscience. Each person is morally entitled to decide for herself what to believe, and not to be brainwashed into thinking what others want her to think.

If I am wrong about that, then so much the worse for the argument, since it obviously relies on the truth of premise 2. But the argument is a failure even if 2 is true. For premise 1 is false.

From the fact that we each have a right to our opinions, nothing at all follows about their plausibility. I was once walking through a forest with a friend who knows a lot about trees. (I don't.) I suggested that the one I was looking at was an ash. It wasn't. He knew it was a larch. Our views were not equally plausible, even though I had as much a right to my opinion as he did his.

There are countless examples of cases in which people have an equal right to an opinion—that is, an equal right not to be forced to change their mind—even though their views are mistaken. Some historical claims are true and others false, even though we each have an equal right to our historical opinions. The same thing can be said of our opinions concerning

economics, trigonometry, basketball strategy, or beer brewing. Most people know more than I do about each of these things, and so my views on these subjects are far less plausible than theirs. And yet my right to hold the views I do is just as strong as anyone else's.

The first premise of the argument confuses two entirely separate matters: whether a person has a right to an opinion, and whether that opinion has any merit. This confusion undermines premise 1, and with it, the argument itself.

4. Moral Objectivity Supports Dogmatism

Pick any blowhard, tyrant, or political fanatic, and there is one thing they all share. They are all ethical objectivists. These are the folks who believe in moral truth with a capital *T*. Luckily for them, they have managed to discover that Truth. All they are trying to do is to let you in on some of it. This may take some shouting, perhaps some coercion, maybe even some killing, but Truth can be pretty demanding.

This thoroughly unpleasant picture yields the following *Argument from Dogmatism*:

1. If there are objective moral standards, then this makes **dogmatism** acceptable.
2. Dogmatism is unacceptable.
3. Therefore, there are no objective moral standards.

Dogmatism is the character trait of being closed-minded and unreasonably confident in one's own opinions. Dogmatism is a vice, and if a theory recommends that we always close our minds to competing ideas, then that theory is very implausible. So premise 2 looks good. But ethical objectivism does not encourage a dogmatic attitude. The first premise of this argument is false.

By itself, the claim that there are objective moral standards is perfectly neutral about how broad-minded we should be. Ethical objectivism is a view about the status of moral claims. It does not tell us what is and is not morally acceptable. All it says is that the correct moral code, *whatever it happens to be*, is objectively true.

But we can say more. If moral truth is not of our own making, then it will not always be easy to discover. And that fact should encourage us to be humble, rather than arrogant and closed-minded. The proper outlook of astronomers and geologists and chemists is that of wonder, a recognition

of one's intellectual limitations, and an appreciation that no matter how smart you are, you'll never know the entire truth about your subject matter. These are appropriate attitudes precisely because there are objective truths in these subjects. Scientists do not get to have the final word about the nature of reality. They might always be corrected by a later generation of thinkers.

If ethics, too, is a subject whose truths are objective, then we should also be open-minded about moral matters. It's perfectly consistent to say that the answers to some questions are objectively true, even though you're not sure what those answers are. If, in ethics, our say-so doesn't make it so, then we are always liable to error. That should give us pause, and it should alert us to the dangers of being dogmatic.

Further, if each person does get to have the final word about morality—if an act is right just because a person approves of it—then people will almost never be morally mistaken. Moral knowledge would be extremely easy to come by. All you'd have to do is to check to see how you feel about an action. If you like it, it's right; dislike it, wrong. If each person is the measure of morality, then we are practically infallible about moral matters. And *that* seems to be a perfect recipe for dogmatism. For why should I change my mind, or think myself mistaken, if the chances of error are almost zero?

It is true that the worst fanatics among us are always ethical objectivists. But that is not a strike against the theory. Rather, it is a strike against the individuals who misapply it and fail to appreciate the complexity of morality and the much greater possibility of error that an objective morality allows. Ethical objectivism is not committed to saying that moral wisdom is easy to get. In fact, as we have seen, objectivism makes such wisdom harder to come by than its competitors do. And so objectivism does not license dogmatism. Thus the first premise of this argument is mistaken. The argument is therefore unsound.

5. Moral Objectivity Supports Intolerance

A very popular reason for rejecting ethical objectivism is a concern for tolerance. People in open societies rightly value tolerance, but many think that tolerance would be threatened if moral standards were objectively correct. If some moral codes are better than others, then what's to stop those with the upper hand from lording it over those who embrace a faulty code of conduct?

Indeed, these critics say that the best way to support tolerance is to assume that all moral views are as good as any other. For if that were so, then no one would be in a position to suppress the lifestyles of those who march to the beat of a different drummer. We would have to agree to disagree, since no one's moral outlook would be better than anyone else's. That is what is needed to support tolerance.

We can trace this line of thinking in the *Argument from Tolerance*:

1. Tolerance is valuable only if the moral views of different people are equally plausible.
2. If ethical objectivism is true, then the moral views of different people are not equally plausible.
3. Therefore, if ethical objectivism is true, then tolerance is not valuable.

That second premise is true. Ethical objectivism rejects the idea of moral equivalence. Some moral views are better than others.

But the first premise is false. In fact, ethical objectivism is much better than moral skepticism at supporting tolerance. The basic reason is this: *If all moral views are equivalent, then a tolerant outlook is no better than an intolerant one.* The outlook of a committed bigot would be as plausible as yours or mine.

Indeed, we can easily frame a counterargument that shows why the value of tolerance poses a threat to skepticism, rather than to objectivism:

1. If all moral views are equally plausible, then moral views supporting tolerance and those supporting intolerance are equally plausible.
2. These moral views are not equally plausible.
3. Therefore, some moral views are less plausible than others.

The first premise must be correct. And those who value tolerance will want to embrace the second. The conclusion follows directly.

Those who favor tolerance tend to regard its value as universal—good for everyone and every society. This applies especially to areas plagued by intolerance, since tolerance is needed most just where it is least enjoyed. Yet if individuals have the final word on what is morally right, then those who are *fundamentally* intolerant—intolerant at their core, in their deepest beliefs—are making no mistake. The same goes for societies. If social codes, rather than individuals, are the measure of morality, then deeply intolerant societies are no worse than freer ones. Their rejection of tolerance is as plausible as your endorsement of it. That should be little comfort to those who value tolerance.

Those who think of tolerance as very valuable will want to say that tolerance is morally required even for those people and those societies that despise it. Such a view is perfectly compatible with ethical objectivism.

6. Moral Disagreement Undermines Moral Objectivity

A classic argument against moral objectivity takes its cue from a simple observation: there is a lot more disagreement in ethics than there is in science. And there is a ready explanation for this. Scientists are trying to understand the nature of objective reality, whereas in ethics, there is no objective reality to be discovered. When it comes to morality, we are merely expressing our personal opinions, ones that have been obviously shaped by the time and place in which we've been raised. Different upbringings, different moral outlooks. But scientists the world over can agree on a wide set of truths, no matter their religious or cultural backgrounds.

The *Argument from Disagreement* nicely summarizes this line of thought:

1. If well-informed, open-minded, rational people persistently disagree about some claim, then that claim cannot be objectively true.
2. Well-informed, open-minded, rational people persistently disagree about all moral claims.
3. Therefore, no moral claim can be objectively true.

Perhaps premise 2 is too strong. *Maybe* there are some moral claims that every smart, rational, open-minded person accepts. But without a lot more investigation, it would be premature to assume that this is so.

What is clearly true is that for any moral claim—even one you find to be just obvious—there will always be someone else who thinks that it is false. But that doesn't show that premise 2 is true, since such people may not be well informed, or open-minded, or rational. Indeed, moral disagreement might well be a product of sloppy reasoning, of not having enough facts under our belt, of having a personal stake in the outcome, or of a general prejudice. What if we were able to correct for these sources of error? Imagine people who were absolutely on top of *all* of the details, say, of affirmative action policies, who were free of personal bias and other prejudices, and who were able to reason flawlessly. Perhaps they'd all agree about whether affirmative action is morally acceptable.

Perhaps. But I share the skeptic's concerns here, and am not sure that even perfectly ideal reasoners would agree about every moral issue. So let's accept, at least for the moment, that premise 2 is true. What of premise 1?

That premise must be false. There are counterexamples galore. Brilliant physicists disagree about whether the fundamental elements of matter are subatomic strings; eminent archaeologists disagree about how to interpret the remains discovered at ancient sites; the finest philosophers continue to debate whether God exists. And yet there are objective truths in each area. There are objective truths about the fundamental nature of the physical world, about the nature of various prehistoric tribes, about whether there is or isn't a God. We may never know these truths, but our opinions on these matters must answer to an objective reality. Our views don't make physical or archaeological or philosophical claims true; the facts are what they are, independently of what we think of them.

There is another reason to doubt premise 1: The first premise is itself the subject of deep disagreement. Really smart people still argue about whether it is true. And so, if such disagreement is enough to undermine objective truth, then the premise, by its own lights, can't be objectively true! And it certainly isn't "relatively" true—true just because I, or my society, believe in it. The premise, then, is false.

So deep disagreement, even among the best minds, is not enough to show that skepticism in an area is correct. As a result, the many disagreements we see in ethics are perfectly compatible with the objectivity of ethics.

7. Atheism Undermines Moral Objectivity

Recall (from chapter 5) the famous thought of Ivan Karamazov, one of Dostoevsky's finest creations: If God is dead, then everything is permitted. His guiding thought is that true morality can exist only if God underwrites its authority. Morality is a sham if God does not exist, because the only way morality could rest on solid foundations is by being authored by God.

Some atheists have taken up this line of thinking and have used it to justify moral skepticism. If they are right, and God does not exist, then morality can't possibly be objective. The *Argument from Atheism* expresses this outlook:

1. Morality can be objective only if God exists.
2. God does not exist.
3. Therefore, morality cannot be objective.

I'm going to make things much easier on myself by leaving that second premise alone. If it's false, and God exists, then the argument crumbles. But let's just assume for now that there is no God. Then what?

Well, if premise 1 is true, and objective morality really does depend on God, then moral skepticism is vindicated. Many people think that 1 is true. They reason as follows. Moral laws, like other laws, must have an author. But if the laws are objective, then (by definition) no human being can be their author. So who is? Three guesses.

This reasoning has always been very popular.⁴ But it is mistaken. It rests on this key assumption: *Laws require lawmakers*. Suppose this assumption is true. It then follows that objective laws need lawmakers, too. But human beings cannot play this role, since objective truths are true independently of human opinion. That leaves only God to do the work.

But if atheism is true, then the crucial assumption is false. Laws would not require lawmakers. Atheists believe that there are objective laws—of logic, physics, genetics, statistics, etc. And yet if God does not exist, these laws have no author. We discovered these laws. We invented the words to describe the laws. But they are not true because we believe them to be. Their truth is objective, not subjective. If atheists are correct, no one authored such laws.

Thus if atheism is true, objective laws do *not* require lawmakers. So, for all we know, objective moral laws do not require a lawmaker, either. Atheists might say, though, that *moral* laws require lawmakers, even though other laws do not. But why single out morality like that? Surely we'd need an excellent reason for thinking that most objective laws need no author, though moral ones do.

Perhaps there is an explanation of this difference. But we can't just assume there is. For atheists who think so, they must point to something special about morality that requires its laws to have an author, even though all other objective laws lack one.

Until they do, we must think that the Argument from Atheism is unpersuasive. It will obviously do nothing to convince religious believers, since it just assumes (in premise 2) that they are wrong. But even if atheists are correct, and God does not exist, premise 1 is highly doubtful, because its best support is flawed. That support comes from the assumption that laws require lawmakers—an assumption that atheists themselves should not accept.

8. The Absence of Categorical Reasons Undermines Moral Objectivity

Most people think that all moral duties come prepackaged with a special power. They automatically supply people with reasons to obey them. And it doesn't matter what we care about. If it's really your duty to repay that loan or help your aged grandparents, then you've got an excellent reason to do so—even if doing these things fulfills none of your desires.

That's unusual. My reasons for writing this book, using my treadmill, getting a regular tune-up for my car, all depend on what matters to me. Most reasons are like this. The reasons that come from morality, however, are categorical. They apply to us regardless of what we care about.⁵

Many philosophers cannot see how categorical reasons are possible. Their puzzlement has given rise to a powerful *Argument from Categorical Reasons* against ethical objectivism:

1. If there are objective moral duties, then there are categorical reasons to obey them.
2. There are no categorical reasons.
3. Therefore, there are no objective moral duties.

This argument has convinced some very smart philosophers. And they may be right to be convinced. But for those with objectivist leanings, there are two lines of response. Since the argument is logically perfect, objectivists have to reject either premise 1 or premise 2.

Some challenge premise 1. They deny that objective moral duties must supply us with reasons for action. It may be that some people have no reason to do what morality requires of them. Whether there are objective moral standards is one thing; whether they supply us with reasons to obey them is another. The answer to the first question may be *yes*, even if the answer to the second is a disappointing *no*. If this line of thinking is right, then we will have to abandon the age-old hope of showing that everyone has reason to be moral.⁶

The second strategy stands by the first premise, but rejects the second. On such a view, objective moral duties really do provide categorical reasons—and these reasons exist. There are reasons to behave in certain ways, even if such behavior doesn't benefit us or satisfy any of our desires. To make this reply a success, we must reject the most popular view of reasons, which says that you have a reason to do something only if doing it will promote self-interest or get you what you care about.

The best way to criticize this popular view is by example. (Recall the discussion in chapter 8, p. 108.) Suppose you are hiking along a cliff path and notice a stranger who is absent-mindedly walking from the opposite direction. You see that he's about to take a wrong step and plunge to his death. There is a reason to yell to him and alert him of the danger. And that reason applies to you even if you don't care a bit about the stranger or about the pats on the back you'll receive when the story gets out. There is something to be said on behalf of your warning him, something that favors it, that justifies it, that makes it a legitimate thing to do. These are just different ways of saying the same thing: There is a reason for you to save that stranger's life, even if doing so won't make you any better off or get you anything you care about.

In short, those who believe that morality is objective must show either that its duties do not have to supply categorical reasons, or argue that such reasons exist. Many philosophers nowadays are developing these strategies, though just as many others (surprise!) are working to ensure that they do not succeed.

9. Moral Motivation Undermines Moral Objectivity

Ask yourself this question: If you sincerely judge an action to be your duty, aren't you automatically motivated (at least a little bit) to do it? If you think a plan or a policy is a morally good one, aren't you moved to some extent to help it along? If you answered yes, then you share the belief that moral judgments are motivational by their very nature. Their essence is to move people to act.

It's not just a coincidence or some kind of minor miracle that moral judgments so reliably move us to act. They have this power because, at their core, what they do is express the very things that cause us to act—our desires, cares, commitments, and emotions. When we judge an action wrong, we are expressing our distaste or hatred of it, our desire that it not be done, our concern for those who might be harmed by it. Our moral judgments express our feelings, and our feelings are our basic motivations. That is why our moral judgments are so easily able to get us to act.

Contrast this picture with another one. Beliefs, unlike moral judgments, are not really in the business of getting us to act in certain ways. They are focused on stating the facts, on reporting the truth, on describing reality. If I believe that there is a computer monitor in front of me, two cats nearby, and an oriental rug underfoot, I am not moved to act in any way at all. If I want to use the Internet, or pet my cats, or vacuum the carpet, then

these beliefs will help to direct my actions. But the key here is that beliefs can do this only by attaching themselves to my desires. If I didn't want to use the Internet, pet my cats, or vacuum my rug, then none of these beliefs would have helped at all in guiding my actions.

This contrast between moral judgment, on the one hand, and belief, on the other, inspired David Hume to construct the following *Motivational Argument*. Generations of moral skeptics have found it compelling:

1. Moral judgments are able, all by themselves, to motivate those who make them.
2. Beliefs are never able, all by themselves, to motivate those who hold them.
3. Therefore, moral judgments are not beliefs.
4. If moral judgments are not beliefs, then they can't be true.
5. Therefore, moral judgments can't be true.

Have a look at that conclusion. Unlike the previous eight arguments, this one doesn't say explicitly that ethical objectivism is false. But that will be cold comfort to the objectivist. If the conclusion of this argument is correct, then moral judgments can't be true. And if they can't be true, then they can't be objectively true. And if they can't be objectively true, then ethical objectivism is false.

Of the three premises in this argument, number 4 is pretty secure. For if moral judgments are not beliefs, then they are expressions of plans, orders, commitments, desires, or emotions. Such expressions are not true or false.⁷ Suppose that I say that genocide is evil. And suppose that I am not thereby stating a belief, but rather expressing my emotions or commitments. What I'm really saying is "Don't commit genocide!" or "Genocide—grrrr." These statements aren't true (or false). So if moral judgments are not beliefs, then they can't be true.

That leaves only two ways for objectivists to fight back against this argument. They can try to undermine the first premise or the second. Unsurprisingly, objectivists have done both.

Some objectivists accept 2, and so criticize 1. They say that moral judgments, like all beliefs, need an additional desire in order to move people to action.

They reason as follows: It's possible that sincere moral judgments leave us entirely cold.⁸ True, since most people have moral concerns and want (at least a little) to be moral, moral judgments will motivate most of the people who make them. But some people just don't care about morality.

They judge things right or wrong and yet are completely unmoved. And that shows that premise 1 is false. Moral judgments cannot move people all by themselves. People must want something in order to be motivated. Thus if moral judgments are going to motivate, they must be joined by a desire to be moral, or a desire for something else that morality can help them achieve.

Some objectivists take a different tack. They insist that beliefs alone *can* motivate people to act. But clearly, not just any belief could do that. My belief that three and three are six, or that Peru is in South America, won't move me to do a thing. But **evaluative beliefs**—beliefs that tell us what is good and bad, or right and wrong—may be able, all by themselves, to get us to act. If they can, then premise 2 is false.

Kant was one of those who rejected premise 2. Recall his claim (in chapter 12, pp. 164–5) that the good will involves only our reason, and not our desires or emotions. Reason tells us that something is our duty, and on the basis of that belief alone, we are motivated to do it. As Kant admitted, it isn't clear whether anyone has ever really acted from a good will, and so acted without the aid of any desires.

These issues are still very much at the center of discussion among philosophers. Objectivists will be able to defeat the Motivational Argument only if they can show that (a) moral beliefs can motivate all by themselves, or (b) they can't, but that this is okay, since not all moral judgments end up motivating people anyway.

10. Values Have No Place in a Scientific World

One of the tools that philosophers use when choosing between competing theories is called **Occam's razor**, after the medieval logician William of Occam (1285–1349). Occam's razor tells us never to multiply entities beyond necessity. What this means in practice is simple. When trying to separate fact from fiction, consider something to be real only if you have to assume its existence in order to explain your experiences.

Occam's razor explains why we shouldn't believe in such things as ghosts. Anything they might account for—spooky feelings in graveyards, creaking noises in old houses—can be better explained without assuming that ghosts really exist. Ghosts aren't needed to make sense of what we see and hear. So Occam's razor tells us that they don't exist.

Many people think that objective values are just like ghosts—creatures of our imagination. These critics deny that we really need to rely on moral features in order to explain the way the world works.

Science is our path to understanding the nature of reality. And scientists never have to include moral features in their explanations of molecular structure, biological adaptations, heat transfer—or anything else. Calling something moral or immoral seems like a kind of luxury, one that contributes nothing essential to understanding the ultimate nature of reality.

We can summarize this line of thinking in the *Argument from the Scientific Test of Reality*:

1. If science cannot verify the existence of X, then the best evidence tells us that X does not exist.
2. Science cannot verify the existence of objective moral values.
3. Therefore, the best evidence tells us that objective moral values do not exist.

This argument reflects a basic commitment to the idea that the supernatural does not exist, and that everything in the world can ultimately be explained by science. Since scientific investigation does not tell us whether actions are moral or immoral, good or evil, this seems to leave objective morality out in the cold.

Ethical objectivists have offered two replies to this argument. Some reject premise 1, and its claim that science is the ultimate test of reality. Others accept this test, but argue that moral features can pass it.

The best strategy for criticizing premise 1 is to see that moral features are **normative** ones. Normative features are those that tell us how things *ought* to be, or how we *should* behave. They rely on **norms**: standards of behavior that supply us with ideals or requirements.

The basic idea behind rejecting premise 1 is this. Science tells us how things really are. Science does not tell us how things ought to be. Science describes; morality prescribes. Science has its limits. It is out of its depth when trying to tell us about our ultimate purpose, the goals we ought to aim for, the standards we should live by. Science can tell us a lot. But it can't tell us everything.

Those who think this way are going to see premise 1 as question-begging. On this line of thinking, only those who already reject objective values will agree with this premise. If you think that science is the only way to gain insight into reality, then you've already give up on objective values.

Further, there is some reason to deny that science really does have the final word on *everything*. Consider:

- (T) A claim is true only if science can verify it.

T can't be true. For science cannot verify it. T is not a scientific statement. We cannot test the truth of T by analyzing what we see, hear, taste, feel, or smell. We cannot mathematically test it. There are no lab experiments that will confirm it.

Since T is false, it follows that there are some truths that science cannot confirm. Perhaps moral ones are among them.

Now consider this principle:

(B) You are justified in believing a claim only if science can confirm it.

B is also problematic, since science cannot confirm it. Only philosophy can do that. If we take B at face value, then by its own lights we cannot be justified in thinking that it is true. So we are not justified in thinking that science is the source of all truths.

Despite these criticisms of the reasoning behind premise 1, some objectivists continue to think that science really is the ultimate source of all of our knowledge, and that we should mistrust any claim that science cannot verify. Such objectivists must therefore find a place for moral values within a scientific world. They do this by arguing that moral features are nothing other than ordinary, run-of-the-mill qualities that fit comfortably within a scientific worldview. This kind of view is called **moral naturalism**.

Moral naturalists could claim, for instance, that being morally right is nothing other than maximizing happiness, or that being good is the very same thing as being desired. We can empirically check whether happiness is maximized, or whether people really desire things. On this view, moral features are nothing but a special class of scientific features. There isn't anything mysterious about them. Moral naturalists thus reject premise 2 of the argument, because they think that moral features just are natural (that is, not supernatural) features of our world.

If that is right, then moral values will need to pass the Occam's razor test of reality. Recall that this test tells us that we have reason to believe that things exist only if they help us to explain what we experience. That's why we no longer believe in ghosts (or the tooth fairy or Superman). We have reason to think that apples are real because the taste and feel we experience when we bite into one demand that we acknowledge their existence. I have reason to think that I have ten fingers and ten toes because I see them, I feel them, other people can confirm their existence, etc.

But are moral features really needed to explain our experiences? Naturalists answer that they are. They think, for instance, that Stalin's *evil nature* is what explains his sending millions to their death in prison camps. The

wrongness of slavery explains why slaves and abolitionists opposed it. The *injustice* and *cruelty* of child abuse explain our anger when we learn of it, the child's resentment, and our efforts to protect children from it.

Moral naturalists think that moral features *just are* natural features. We use a different vocabulary to refer to them—we talk of good and evil, right and wrong, rather than neutrinos and quarks and molecules and proteins. But the words we use are not important. What is important is that, as moral naturalists see things, the natural world is the only world there is, and so moral features must be part of that world if they are to exist.

If moral features are natural features, then, like all natural features, science can test for their existence. And, as we have seen, scientists rely on Occam's razor to help determine what really does and does not exist. If moral features are natural ones, then science will indeed confirm that things are right or wrong, good or evil. Moral features will play an essential role in explaining how the world works. Premise 2 would therefore be false.

I have spent the most time on this argument because, among all ten arguments considered here, I think that it presents the most serious threat to ethical objectivism. Philosophers nowadays are involved in major research projects in this area, both to try to sharpen this argument against objectivism and also to develop replies along the lines I have indicated.

Conclusion

Our discussion of these ten arguments has not revealed a single one that confirms the existence of objective moral values. That was deliberate. I had given over the previous two chapters to the critics of objectivism, and showed how each of the anti-objectivist theories encounters some serious problems.

It's only fair, then, that we devote some time to the many worries that people have about how morality could possibly be objective. I have tried here to outline the ones that seem to me to be either very popular or very threatening. The most popular ones actually seem easiest to handle. The less well known arguments strike me as more difficult to rebut.

There is no quick, knock-down argument that will demolish ethical objectivism. Nor is there any short and sweet proof of its truth. I have offered these arguments, and their replies, not in order to create the impression of a victory for either side, but rather to give you a sense of how complicated things can get in this area of philosophy. Those who act as if moral skepticism were obviously true, or just plainly false, have simply gotten it wrong. Matters here, as elsewhere in ethics, are too challenging to admit of pat and easy solutions.

Notes

1. See chapter 15 for much more detailed discussion of this question.
2. See chapter 16 for a presentation of Ross's views.
3. Unlike almost all of the other arguments we've seen in this book, this one has only a single premise: 1. But that's not a problem; the logic of this argument is perfect. If 1 is true, then 2 must be true.
4. For much more on this line of reasoning, see the extended discussion in chapter 5, pp. 61–66.
5. For detailed discussion of categorical reasons, see chapter 8, pp. 102–105 and chapter 11, pp. 152–155.
6. If this line can be defended, then we also have an adequate reply to error theorists. They claim that moral thinking assumes the existence of categorical reasons, but that no such reasons exist. But if morality does not make that assumption, then it may be in good shape even if there are no categorical reasons.
7. Recall the lesson about such expressions from the previous chapter's discussion of expressivism, pp. 299–300.
8. These people are known as *amoralists*. See the discussion in the previous chapter, p. 302, and in chapter 14, pp. 191–195, for more on amoralism.

REFERENCES

..... ☛

- Aquinas, Thomas. *Summa Theologica*. Available in many translations.
- Aristotle. *Nicomachean Ethics*. Available in many translations.
- . *Politics*. Available in many translations.
- Bellow, Saul. *Humboldt's Gift* (New York: Viking, 1975).
- Bentham, Jeremy. *Introduction to the Principles of Morals and Legislation* (1781). Available from many publishers.
- Colfax, Richard. *Evidence Against the Views of the Abolitionists, Consisting of Physical and Moral Proofs, of the Natural Inferiority of the Negroes* (New York: James T. M. Bleakley Publishers, 1833).
- Doctorow, E. L. *The March* (New York: Random House, 2004).
- Dostoevsky, Fyodor. *The Brothers Karamazov* (1880). Available in many translations.
- Fireside, Harvey. *Soviet Psychoprisoners* (New York: W.W. Norton, 1982).
- Foot, Philippa. *Natural Goodness* (New York: Oxford University Press, 2001).
- Fraser, Flora. *Pauline Bonaparte: Venus of Empire* (New York: Alfred A. Knopf, 2009).
- Gilligan, Carol. *In a Different Voice* (Cambridge, Mass.: Harvard University Press, 1982).
- Golding, William. *The Lord of the Flies* (1954). Available from many publishers.
- Grayson, William John. *The Hireling and the Slave* (Charleston, S.C.: John Russell, 1855).
- Hare, R. M. "What Is Wrong with Slavery," *Philosophy and Public Affairs* 8 (1979): 103-21.
- Hobbes, Thomas. *Leviathan* (1651). Available from many publishers.
- Hornby, Nick. *How to be Good* (New York: Riverhead Books, 2001).
- Hume, David. *A Treatise of Human Nature* (1739). Available from many publishers.
- Huxley, Aldous. *Brave New World* (1932). Available from many publishers.

R-2 REFERENCES

- Kant, Immanuel. *Groundwork of the Metaphysics of Morals* (1785). Available in many translations.
- . *Observations on the Feeling of the Beautiful and the Sublime* (1764). Available in many translations.
- Karr, Mary. *The Liars' Club* (New York: Penguin, 1995).
- Kesey, Ken. *One Flew over the Cuckoo's Nest* (New York: Viking Press, 1962).
- Levi, Primo. *The Drowned and the Saved* (New York: Knopf, 1986).
- Mill, John Stuart. *Utilitarianism* (1861). Available from many publishers.
- . *On the Subjection of Women* (1869). Available from many publishers.
- Moore, G. E. *Principia Ethica* (Cambridge, UK: Cambridge University Press, 1903).
- Nozick, Robert. *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- Plato. *Euthyphro*. Available in many translations.
- . *Republic*. Available in many translations.
- Reddaway, Peter, and Sidney Bloch. *Soviet Psychiatric Abuse: The Shadow over World Psychiatry* (Boulder, Colo.: Westview Press, 1984).
- Ross, W. D. *The Right and the Good* (Oxford, UK: Oxford University Press, 1930).
- Rousseau, Jean-Jacques. *Emile: On Education* (1762). Available in many translations.
- Styron, William. *Sophie's Choice* (New York: Random House, 1979).
- Waugh, Alexander. *Fathers and Sons* (New York: Doubleday, 2007).
- Wisconsin v. Yoder* 406 U.S. 205 (1972).

SUGGESTIONS FOR FURTHER READING



Where to Start

Most of the topics covered in these pages are also represented in the companion volume to this book, *The Ethical Life: Fundamental Readings in Ethics and Moral Problems* (Oxford University Press, 2010). That book provides selections of original work by other philosophers. I chose the pieces there with an eye to the introductory student, so most of the material should be fairly accessible to those just beginning their philosophical studies. If you want a relatively short collection that ranges over the main issues discussed here, plus a lot of coverage of specific moral problems, such as abortion, the death penalty, and animal rights, then that might be a good place to start.*

An excellent source for the entire range of philosophical issues, not just those in ethics, is the *Stanford Encyclopedia of Philosophy*, a free online resource containing articles written by experts in the field. The articles are usually pitched to those with little prior knowledge of the topic under discussion: <http://plato.stanford.edu/>.

There are a number of other texts designed to introduce students to the field of moral philosophy. Among the better ones are: James Rachels and Stuart Rachels, *The Elements of Moral Philosophy* 6th ed. (McGraw-Hill 2009); Mark Timmons, *Moral Theory* (Rowman and Littlefield, 2001); and Julia Driver, *Ethics: The Fundamentals* (Blackwell, 2006). Of these, the Rachels and Rachels book is best suited for those with no philosophy background. The Driver and the Timmons books are a bit more advanced.

Three very good anthologies provide fairly accessible survey articles of the major theories in and about ethics. One of these is *A Companion to Ethics* (Blackwell, 1991), edited by Peter Singer. Another is *The Blackwell Guide to Ethical Theory* (Blackwell, 2000), edited by Hugh LaFollette. The last is *The Oxford Handbook of Ethical Theory* (Oxford University Press, 2007), edited by David Copp. *The Encyclopedia of Ethics*

* All sources that appear in **boldface** in this section are included in *The Ethical Life*.

(Garland, 1992), edited by Lawrence and Charlotte Becker, is available in most college and university libraries and offers entries on all of the topics covered in this book.

Hedonism

Epicurus's works are available in many editions. A reliable and well-priced version is *The Epicurus Reader*, edited by L. Gerson and B. Inwood (Hackett, 1994). His *Letter to Menoeceus*, included in that collection, summarizes the main doctrines of his philosophy. W. D. Ross's two-worlds objection to hedonism can be found in *The Right and the Good* (Oxford University Press, 1930), chapter 5. **Robert Nozick's experience machine discussion** can be found in his *Anarchy, State, and Utopia* (Basic Books, 1974), pp. 42–45. **John Stuart Mill's version of hedonism** is presented in chapters 2 and 4 of *Utilitarianism* (many publishers). Jeremy Bentham's version of hedonism can be found in his *Introduction to the Principles of Morals and Legislation* (1781), available from many publishers. Perhaps the most sophisticated contemporary defense of hedonism is offered by Fred Feldman in his very clearly written *Pleasure and the Good Life* (Oxford University Press, 2006). A defense of the view that informed and autonomous happiness is the key to a good life is given by L. W. Sumner in his *Welfare, Happiness, and Ethics* (Oxford University Press, 1995). His book also provides a nice overview of the issues surrounding the nature of the good life.

A very accessible, engaging work for introductory students is Joel Kupperman's *Six Myths About the Good Life* (Hackett Publishers, 2006), which covers hedonism, the desire theory, and other options not discussed here. Those who want more in the way of short selections from classic texts in this area might consult *The Good Life*, edited by Charles Guignon (Hackett Publishers, 1999). On hedonism and happiness more generally, see Nicholas White's historical survey *A Brief History of Happiness* (Blackwell, 2006), and Steven Cahn's anthology *Happiness: Classic and Contemporary Readings in Philosophy* (Oxford University Press, 2007).

Getting What You Want

Very few philosophers have defended the view that satisfaction of our actual desires, based as they often are on ignorance, prejudice, and faulty reasoning, serve as the key to a good life. Contemporary philosophers who come close are Mark Murphy, "The Simple Desire-Fulfillment Theory," *Nous* 33 (1999): 247–72, and Simon Keller's accessible and enjoyable "Welfare and the Achievement of Goals," *Philosophical Studies* 121 (2004): 27–41. **Richard Taylor's "The Meaning of Life,"** from his book *Good and Evil* (Prometheus Books, 2000), pp. 256–68, takes the view that desire satisfaction is what gives a life its meaning.

James Griffin's *Well-Being* (Oxford University Press, 1985), part 1, provides a good discussion of the various difficulties surrounding the desire satisfaction theory, but also offers a qualified defense of the view. The view that the satisfaction of our filtered, more informed desires is the basis of personal welfare is defended by John Rawls, *A Theory of Justice* (Harvard University Press, 1971), pp. 417ff.; Richard Brandt in *A Theory of*

the Good and the Right (Oxford University Press, 1979), pp. 126–29; and Peter Railton, “Facts and Values,” included in his collection of important essays, *Facts, Values, and Norms* (Cambridge University Press, 2003).

A lovely critical discussion of the desire view, with lots of examples meant to damage it and to provide indirect support for his own more Aristotelian view, can be found in Richard Kraut’s *What Is Good and Why?* (Harvard University Press, 2007), chapter 2. Another excellent critical discussion, though less accessible, is Connie Rosati’s “Persons, Perspectives, and Full Information Accounts of the Good,” *Ethics* 105 (1995): 296–325. An absolutely delightful book, chock full of real-life stories and interesting examples, is **Jean Kazez’s *The Weight of Things*** (Blackwell, 2006). She defends an objective view about well-being in chapters 5 and 6.

Morality and Religion

Plato’s *Euthyphro* is available in many translations. At about eleven pages, it’s an enjoyable introduction to Plato’s early work. Perhaps your best bet is to get it packaged with four other Platonic dialogues in an excellent, inexpensive translation by G.M.A. Grube and John Cooper, *Five Dialogues* (Hackett, 2001).

Defenses of the divine command theory tend to be fairly complex and difficult. A pretty accessible version is by the late Philip Quinn, in his article on the theory in Hugh LaFollette’s (ed.) *The Blackwell Guide to Ethical Theory* (2000). Robert Adams is another notable defender of the theory. His work is not easy for the beginner, but “A New Divine Command Theory,” *Journal of Religious Ethics* 7 (1979): 66–79, might not be a bad place to start. A more accessible version of this paper is given in an anthology that I have edited, which ranges across most areas of ethics: *Ethical Theory* (Blackwell, 2007).

Though it focuses on many other issues as well, *God? A Debate Between a Christian and an Atheist* (Oxford University Press, 2004), written by William Lane Craig (the Christian) and Walter Sinnott-Armstrong (the atheist), also contains some common lines of defense and criticism surrounding the divine command theory. It is written in a very lively style. Accessible assessments of the divine command theory can be found in most of the introductory books mentioned in the “Where to Start?” section at the beginning of *Selections for Further Reading*. A critical discussion of the divine command theory that is quite easy to read can be found in Erik Weilenberg’s *Value and Virtue in a Godless Universe* (Cambridge University Press, 2005), chapter 2. Kai Nielsen’s *Ethics Without God* (Prometheus, 1990) is a clear treatment of a number of issues regarding religion and morality, written from the perspective of someone who thinks that ethics is self-standing and has no need of religious input.

Natural Law

The attempt to base morality on human nature can be traced in the West all the way to Aristotle. His *Nicomachean Ethics*, especially books I and II, are the place to start. A fine and helpful translation is offered by Terence Irwin (Hackett, 1999, 2nd edition). Medieval philosopher Thomas Aquinas, whose work continues to exercise the largest

influence on Roman Catholic moral theology, is the essential source for thinking about developments of natural law over the past seven hundred years. Aquinas isn't that approachable; you could dip a toe into the water by having a look at Question 94 of the *Prima Secundae* of his *Summa Theologica*. The *Summa* runs to five volumes and over a thousand pages, but this discussion can be found in almost every shorter collection of Aquinas's works. A good book for beginners is *Aquinas: Selected Writings* (Penguin, 1999), edited by Ralph McInerny.

Important contemporary natural lawyers include John Finnis, whose *Natural Law and Natural Rights* (Oxford University Press, 1980) did much to revive this ethical tradition within secular academic circles. A good scholarly history can be found in Knud Haakonssen's *Natural Law and Moral Philosophy: From Grotius to the Scottish Enlightenment* (Cambridge University Press, 1996). **Philippa Foot's** *Natural Goodness* (Oxford University Press, 2001) is a delightfully written book by a very important moral philosopher.

Psychological Egoism

Though it is nowadays the subject of some debate among scholars, it seems that Thomas Hobbes committed himself to psychological egoism in several passages of his masterpiece, *Leviathan*. This work is available from many publishers; if you have an ear for seventeenth-century English, you will love Hobbes's vigorous style. Joseph Butler, an eighteenth-century bishop, produced criticisms of psychological egoism that many still regard as decisive. See his *Fifteen Sermons Preached at the Rolls Chapel*, the relevant portions of which are presented in *Five Sermons* (Hackett, 1983), edited by Stephen Darwall. David Hume, a master stylist himself, also criticized psychological egoism in appendix 2 of his *Enquiry Concerning the Principles of Morals*, available from many publishers.

A very clear, approachable article that explains the motivations and problems of psychological egoism is Joel Feinberg's "Psychological Egoism," in Feinberg and Shafer-Landau, eds., *Reason and Responsibility* (Wadsworth, many editions). Empirical work on psychological egoism is given a careful review by C. D. Batson, *The Altruism Question: Toward a Social-Psychological Answer* (Erlbaum, 1991). Elliot Sober and David Sloan Wilson provide a scientifically well-informed and philosophically sophisticated approach to the merits of psychological egoism in their *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Harvard, 1999).

Ethical Egoism

What I have called "The Best Argument for Ethical Egoism" can be pieced together from claims made by **Thomas Hobbes in his** *Leviathan*. The claim that we have reason to do only what will serve self-interest is defended by David Gauthier in his important (but difficult) book *Morals by Agreement* (Oxford University Press, 1986), chapter 2. A crystal-clear historical survey of this thesis about reasons is given in Robert Shaver's *Rational Egoism* (Cambridge University Press, 1999).

An easy-to-read pair of articles on the merits of ethical egoism is offered by Brian Medlin, "Ultimate Principles and Ethical Egoism," and Jesse Kalin, "On Ethical Egoism," both included in David Gauthier's anthology *Morality and Self-Interest* (Prentice Hall, 1970). Lester Hunt defends the view that ethical egoism will not require us to violate the rules of conventional morality in his "Flourishing Egoism," *Social Philosophy and Policy* 16 (1999): 72–95. Gregory Kavka's "The Reconciliation Project" is a terrific exploration of how far self-interest and conventional morality can be reconciled. His view is a bit less optimistic than Hunt's, but only a bit. It can be found in David Zimmerman and David Copp, *Morality, Reason, and Truth* (Rowman and Allanheld, 1984), pp. 279–319.

Consequentialism

John Stuart Mill's *Utilitarianism* is the place to start. It is short and elegant; many editions are available. Perhaps the greatest utilitarian treatise ever written is Henry Sidgwick's *The Methods of Ethics* (1907; also available from many publishers). Sidgwick's writing style does not endear him to the reader, however—especially the introductory reader. R. M. Hare's writing style, by contrast, is clean and elegant; his sophisticated defense of utilitarianism can be found in his *Moral Thinking* (Oxford University Press, 1981).

A very influential defense of act utilitarianism and critique of rule utilitarianism are given in J.J.C. Smart, "Extreme and Restricted Utilitarianism," *Philosophical Quarterly* 6 (1956): 344–54. A terrific book that set the terms of the debate for the next generation of moral philosophers is one that Smart wrote with Bernard Williams, *Utilitarianism: For and Against* (Cambridge University Press, 1973). Brad Hooker defends rule consequentialism in a clear and accessible way in his contribution to Hugh LaFollette's *The Blackwell Guide to Ethical Theory* (Blackwell, 2000), and in his book *Ideal Code, Real World* (Oxford University Press, 2000).

There are several good collections of articles and book excerpts on the subject of consequentialism. The contents usually reflect work being done by and for fellow philosophers, so the going isn't always easy. Perhaps the one that contains the greatest bang for the buck for the introductory student is Jonathan Glover's *Utilitarianism and Its Critics* (Prentice Hall, 1990). Samuel Scheffler's *Consequentialism and Its Critics* (Oxford University Press, 1988) contains many fine articles, but the going is sometimes quite difficult. Stephen Darwall does a nice job collecting classic readings and important contemporary ones in his *Consequentialism* (Blackwell, 2002).

Kantian Ethics

Kant's writing is not at all easy to work through. The most accessible (or rather, least inaccessible) of his works is also the shortest: *The Groundwork of the Metaphysics of Morals*. It comes in at a bit under sixty pages; parts 1 and 2 (there are three parts in all) can occasionally be read with pleasure and ready comprehension. The best

translation is offered by Mary Gregor, with an excellent introduction by Christine Korsgaard (Cambridge University Press, 1998). The translations of Lewis White Beck and H. G. Paton are also good. Paul Guyer's *Kant's Groundwork of the Metaphysics of Morals: A Reader's Guide* (Continuum, 2007) is a helpful book to have by one's side when reading this classic text. For the intrepid reader who wants more Kant than this, try his *Metaphysics of Morals*, also translated by Mary Gregor (Cambridge University Press, 1996).

A number of fine philosophers have written engaging essays that interpret and apply Kant's moral philosophy and demonstrate its contemporary relevance. You might try Barbara Herman's *The Practice of Moral Judgment* (Harvard University Press, 1993) and *Moral Literacy* (Harvard University Press, 2008), Thomas E. Hill, Jr.'s *Autonomy and Self-Respect* (Cambridge University Press, 1991) and *Dignity and Practical Reason* (Cambridge University Press, 1992), Christine Korsgaard's *Creating the Kingdom of Ends* (Cambridge, 1996), or Onora O'Neill's *Constructions of Reason* (Cambridge University Press, 1990).

On the value of integrity and conscientiousness, see **Jonathan Bennett's wonderful article, "The Conscience of Huckleberry Finn,"** *Philosophy* 49 (1974): 123–34.

Social Contract Theory

Thomas Hobbes's *Leviathan* is the place to start. For those with only a relatively short amount of time on their hands, go directly to **chapters 13–15**, and then keep reading as time permits. John Locke's *Second Treatise of Government* and Jean-Jacques Rousseau's *The Social Contract* (both available in many editions) are also important classics in this tradition. Locke's short book was especially influential in the thinking of the authors of the Declaration of Independence and the U.S. Constitution.

The Hobbesian approach to morality is given an important and sophisticated update by David Gauthier, in his *Morals by Agreement* (Oxford University Press, 1986). Gregory Kavka's *Hobbesian Moral and Political Theory* (Princeton University Press, 1986) is wonderful both as commentary and as good, clear-headed philosophy.

John Rawls's *A Theory of Justice* (Harvard University Press, 1971; revised edition 1999) was recognized as a masterpiece upon its publication. A shorter presentation of his central ideas can be found in *Justice as Fairness: A Restatement* (Harvard University Press, 2001, 2nd ed.). Rawls's theory is, as its title suggests, a theory of justice rather than a theory about the whole of morality. Still, its influence in ethics, as well as in social and political philosophy, would be difficult to overstate.

T. M. Scanlon's very important ethical theory, which he terms "contractualism," is a contemporary offshoot of the social contract theory. He presents it in his book *What We Owe to Others* (Harvard University Press, 1998). It's long and rarely an easy go for the beginner; those who want a briefer introduction to his thinking are advised to have a look at his paper, "Contractualism and Utilitarianism," included in

a collection edited by Amartya Sen and Bernard Williams, *Utilitarianism and Beyond* (Cambridge University Press, 1982), pp. 103–28.

A nice collection of excerpts and essays from social contract theorists is offered by Stephen Darwall, ed., *Contractarianism/Contractualism* (Blackwell, 2002).

Ethical Pluralism

A nice collection of articles, some in defense of absolutism and others critical of it, can be found in Joram Haber, ed., *Absolutism and Its Critics* (Rowman and Littlefield, 1994). An unusual introduction to ethics, one that contains a number of pieces that focus on the Doctrine of Double Effect and the Doctrine of Doing and Allowing, is *Ethics: Problems and Principles* (Wadsworth, 1992), edited by John Martin Fischer and Mark Ravizza.

A very interesting piece that renewed interest in the DDE and DDA is **Philippa Foot's "Abortion and the Doctrine of Double Effect,"** *Oxford Review* 5 (1967): 5–15. Foot introduced the now-famous trolley problem to the philosophy literature, as well as numerous other examples that have stimulated philosophical discussion for the past several decades. See also Judith Jarvis Thomson's articles "The Trolley Problem," and "Killing, Letting Die, and the Trolley Problem," both reprinted in her marvelous collection *Rights, Restitution, and Risk* (Harvard University Press, 1986). Thomson introduces a number of important variations on Foot's example and argues for anticonsequentialist principles meant to explain why it is only sometimes, and not always, permitted to minimize harm.

Stephen Darwall's collection, *Deontology* (Blackwell, 2002) contains a number of important papers that explore the idea that certain kinds of actions are intrinsically right or wrong, and discuss the question of whether there are any absolute moral rules.

W. D. Ross presents his ethic of prima facie duties in **chapter 2 of *The Right and the Good*** (Oxford University Press, 1930). An excellent article defending Ross against a variety of criticisms is David McNaughton's "An Unconnected Heap of Duties?" *Philosophical Quarterly* 46 (1996): 433–47. McNaughton was earlier a fan of ethical particularism; his stimulating and well-written introduction to ethics, *Moral Vision* (Blackwell, 1988), offers a defense of particularism. Jonathan Dancy put particularism on the map with a series of articles in the 1980s; his book *Ethics Without Principles* (Oxford University Press, 2004) offers his latest views on the matter. Gerald Dworkin has written a very accessible article on how particularists can gain moral knowledge: "Unprincipled Ethics," *Midwest Studies in Philosophy* 20 (1995): 224–38.

Virtue Ethics

Study of virtue ethics must begin with **Aristotle's *Nicomachean Ethics***. Many good translations are available. In addition to the one by Terence Irwin, mentioned in the "Natural Law" section, the one undertaken by our old friend W. D. Ross, the preeminent Aristotle scholar of his day, is also excellent. It has been updated by

J. O. Urmson and J. L. Ackrill (Oxford University Press, 1998). Christopher Rowe has also provided a fine translation, aided by Sarah Broadie's substantial and illuminating notes, in their edition of the *Nicomachean Ethics* (Oxford University Press, 2002).

The Blackwell Guide to Aristotle's Nicomachean Ethics (Blackwell, 2006), edited by Richard Kraut, is highly recommended. It includes instructive articles on many important aspects of Aristotle's ethical thought by a who's who of leading scholars.

The best short overview of virtue ethics that I have read is by Julia Annas, in her contribution to David Copp's *The Oxford Handbook of Ethical Theory* (Oxford University Press, 2007).

Two excellent collections on virtue ethics are Stephen Darwall, ed., *Virtue Ethics* (Blackwell, 2002), and Michael Slote and Roger Crisp, eds., *Virtue Ethics* (Oxford University Press, 1997).

Alasdair MacIntyre's much-discussed *After Virtue* (University of Notre Dame Press, 1981) rekindled interest in this tradition after a long period of dormancy in the United States and Britain. Other important recent works in virtue ethics include Rosalind Hursthouse's *On Virtue Ethics* (Oxford University Press, 2000), Michael Slote's *Morals from Motives* (Oxford University Press, 2003), and Christine Swanton's *Virtue Ethics: A Pluralistic View* (Oxford University Press, 2005). Martha Nussbaum is a wonderful writer and has done a lot of work on Aristotle and ethics. One of her most important papers defends Aristotle, and virtue ethics, from the charge of relativism. See her "Non-Relative Virtues: An Aristotelian Approach," in *Midwest Studies in Philosophy* 13 (1988): 32–53.

The Ethics of Feminism

A good place to start is Hilde Lindemann's *An Invitation to Feminist Ethics* (McGraw-Hill, 2006), written with non-philosophers and beginning students in mind. Its **first chapter** provides a nice, brief overview of feminist ethics, while chapter 4 offers a succinct review of feminist criticisms of utilitarianism, Kantianism, and contractarianism. But the entire book is worth a read.

Those with an interest in the ethics of care should start with Carol Gilligan's fascinating *In a Different Voice* (Harvard University Press, 1982), and proceed to Nel Noddings's *Caring: A Feminine Approach to Ethics and Moral Education* (University of California Press, 1984). Two recent studies by important philosophers are Michael Slote's *The Ethics of Care and Empathy* (Routledge, 2007) and Virginia Held's *The Ethics of Care: Personal, Political, Global* (Oxford University Press, 2007). Those who want a much briefer, but still substantial treatment of the subject, would do well to have a look at Held's "The Ethics of Care," in *The Oxford Handbook of Ethical Theory* (Oxford University Press, 2007), edited by David Copp.

Helpful overviews of the huge range of work in feminist ethics include Alison Jaggar's "Feminist Ethics: Projects, Problems, Prospects," in Claudia Card's collection *Feminist Ethics* (University Press of Kansas, 1991), pp. 78–103. Another nice overview,

entitled "Feminist Ethics," is written by Rosemarie Tong and Nancy Williams, and appears online in the *Stanford Encyclopedia of Philosophy*.

For a taste of the many moral issues that receive fresh light when seen from a feminist perspective, you might try *Feminist Philosophies* (Prentice-Hall, 1992), edited by Janet Kourany, James Sterba, and Rosemary Tong. Cheshire Calhoun's collection, *Setting the Moral Compass: Essays by Women Philosophers* (Oxford University Press, 2004), includes essays by a roster of outstanding philosophers writing on issues in and around feminist philosophy.

The Status of Morality

Most of the work done in metaethics is not that accessible for beginning students. I have written a very elementary introduction to metaethics, entitled *Whatever Happened to Good and Evil?* (Oxford University Press, 2004), designed for those with no prior philosophy knowledge. Robert Audi's *Moral Value and Human Diversity* (Oxford University Press, 2007) is also pitched to an introductory audience. For a more advanced treatment, Alexander Miller's *An Introduction to Contemporary Metaethics* (Polity, 2004) is a valuable resource. A historically informed survey of views on the topic is given by Stephen Darwall in his *Philosophical Ethics* (Westview, 1997).

The early chapters of Book III of David Hume's *Treatise of Human Nature* have set the terms of the debate in metaethics for the past two and a half centuries. Hume's work has inspired important contemporary philosophers such as Gilbert Harman, whose own introduction to ethics, *The Nature of Morality* (Oxford University Press, 1977) contains **an early chapter that challenges ethical objectivism**. Harman is also the most prominent contemporary moral relativist. His paper "Moral Relativism Defended," *Philosophical Review* 85 (1975): 3–22, is worth seeking out. It and four other interesting essays in defense of relativism are included in his *Explaining Value* (Oxford University Press, 2000).

J. L. Mackie's now-classic defense of the error theory is given in **the first chapter of his *Ethics: Inventing Right and Wrong*** (Penguin, 1977). Australian philosopher Richard Joyce defends the error theory with verve in his *The Myth of Morality* (Cambridge University Press, 2001).

Expressivism entered the scene in the 1930s with chapter 6 of A. J. Ayer's *Language, Truth, and Logic*. This is well worth a read, as Ayer pulls no punches and is a lovely writer. Simon Blackburn's work is the most accessible among contemporary expressivists, though written with fellow philosophers in mind. His *Essays in Quasi-Realism* (Oxford University Press, 1993) contains many important papers. But a better place to start would be his introduction to philosophy, *Think* (Oxford University Press, 1999), chapter 8, which is more accessible and written with his characteristic elegance.

Defense of a kind of ideal-observer view is given by Michael Smith in his important book *The Moral Problem* (Blackwell, 1994). This book provides an excellent way

in to the many problems in metaethics. Smith has also written an introductory article that highlights a number of the themes of his book: "**Realism**," in Peter Singer, ed., *Ethics* (Oxford: Oxford University Press, 1994), pp. 170–76.

For a defense of ethical objectivism that is as clear as philosophical writing gets, see David Brink, *Moral Realism and the Foundations of Ethics* (Cambridge University Press, 1989). Brink is a moral naturalist whose book offers detailed coverage of most of the major issues in metaethics.

For a wide-ranging collection containing many classic and contemporary writings on the subject, along with a dozen substantial introductory essays designed with the student reader in mind, see Terence Cuneo and Russ Shafer-Landau, eds., *Foundations of Ethics: An Anthology* (Blackwell, 2006).

GLOSSARY



Absolute: never permissibly broken; violating an absolute moral rule is always wrong.

Act consequentialism: The normative ethical theory that says that an act is morally right just because it produces the best actual or expected results.

Act utilitarianism: The version of act consequentialism that says that only well-being is intrinsically valuable, and so says that an act is morally right just because it maximizes overall well-being.

Ad hominem attack: An attempt to undermine the position of an opponent by criticizing his motives or character.

Agnostics: Those who suspend judgment on the question of whether God exists.

Altruism: The direct care and concern to improve the well-being of someone other than yourself.

Ambiguous: having two or more meanings.

Amoralists: Those who do not care about living up to the moral views they sincerely hold.

Argument: Any chain of thought in which premises are enlisted in support of a particular conclusion.

Atheism: The belief that God does not exist.

Autonomy: The capacity to determine for yourself the principles that you will live by. It can also refer to your ability to live according to your own plan of life.

Begging the question: Arguing on the basis of a reason that will appeal only to people who already accept the argument's conclusion.

Categorical imperative: A command of reason that requires a person's obedience regardless of whether such obedience gets him anything he wants.

Categorical reason: A reason to do something that applies to a person regardless of her desires.

- Coherentism:** The view that we are justified in believing a claim to the extent that it supports, and is supported by, other beliefs we hold.
- Conceptual truth:** A true claim that can be known just by understanding it. Such a claim is true just by virtue of the concepts it contains—that's why understanding it enables one to know it. An example: bachelors are unmarried men.
- Consent, tacit:** See **tacit consent**.
- Consequentialism:** A family of normative ethical theories that share the idea that the morality of actions, policies, motives, or rules depends on their producing the best actual or expected results. See also: **act consequentialism, rule consequentialism, act utilitarianism, rule utilitarianism**.
- Continent:** doing the right thing while suppressing desires that tempt one away from doing one's duty.
- Contractarianism:** See **social contract theory**.
- Cultural relativism:** The view that an act is morally right just because it is allowed by the guiding ideals of the society in which it is performed, and immoral just because it is forbidden by those ideals.
- Decision procedure:** Any method designed to guide us in successfully deliberating about what to do.
- Deist:** One who believes that God exists, created the universe, and then refrained from becoming involved in human affairs.
- Desire satisfaction theory:** A theory of human well-being that claims that the satisfaction of your actual or informed desires is necessary and sufficient to improve your welfare.
- Doctrine of Doing and Allowing:** The view that it is always morally worse to do harm than to allow that same harm to occur.
- Doctrine of Double Effect:** The view that if your goal is worthwhile, you are sometimes permitted to act in ways that foreseeably cause certain harms, though you must never intend to cause those harms.
- Dogmatism:** The trait of being closed-minded and unreasonably confident of the truth of one's views.
- Empirical truth:** A true claim that can be known only by means of evidence gained through the senses. Understanding what such a claim says is not enough to know whether it is true—you have to check the claim "against the world" to test it. An example: the Empire State Building is 1453 feet tall.
- Error theory:** The metaethical view that there are no moral features in this world; no moral judgments are true; our sincere moral judgments try, and always fail, to describe the moral features of things; and there is no moral knowledge.
- Ethical egoism:** The normative ethical theory that says that actions are morally right just because they maximize self-interest.
- Ethical monism:** The view that there is only one moral rule that is absolute and fundamental.
- Ethical objectivism:** The view that there is at least one objective moral standard.

Ethical particularism: The view that there are neither any absolute nor any *prima facie* moral rules. According to ethical particularism, no feature of the world is always morally relevant, and none is always morally decisive.

Ethical pluralism: The view that there are at least two, and possibly more, fundamental moral rules.

Ethical relativism: The view that correct moral standards are relative to individual or cultural commitments. Ethical relativism can take two forms: **cultural relativism** or **ethical subjectivism**.

Ethical subjectivism: The view that an act is morally right just because (a) I approve of it, or (b) my commitments allow it. An action is wrong just because (a) I disapprove of it, or (b) my commitments forbid it.

Eudaimonia: The state of living well; happiness, or flourishing.

Evaluative beliefs: Beliefs that evaluate something, and so assess it as good or bad, virtuous or vicious, etc.

Exemplar, moral: See **moral exemplar**.

Exemplary punishment: Punishment designed to make an example of the one who is punished.

Expressivism: The version of moral nihilism that denies that there are any moral features in this world; claims that there is nothing for moral judgments to be true of; and analyzes moral judgments as expressions of emotions, orders, or commitments, none of which are the sorts of things that can be true or false.

Fidelity: Being faithful to one's word; keeping one's promises.

Free-rider problem: A situation in which people are able to obtain a share of some common good without contributing to it. In such situations, it appears to be rational (if your withholding can go unnoticed) to refrain from contributing, thus enjoying the good at no expense to yourself. The problem is that if enough people act rationally, then there will not be enough resources to produce the relevant good, thus harming everyone.

Fundamental: A moral rule is fundamental just in case its justification does not depend on any more general or more basic moral rule.

Golden rule: The normative ethical principle that says that your treatment of others is morally acceptable if and only if you would be willing to be treated in exactly the same way.

Good will: The ability to reliably determine what your duty is, and a steady commitment to do your duty for its own sake.

Hedonism: The view that pleasure is the only thing that is intrinsically valuable, and pain (or unhappiness) is the only thing that is intrinsically bad.

Hypothetical imperative: A command of reason that requires a person to take the needed means to getting what she wants.

Iconoclasts: People whose views differ radically from the conventional wisdom of their society.

Ideal observers: Those (probably imaginary) people who are fully informed, perfectly rational, and otherwise perfectly suited to determine the content of morality.

Imperative, categorical: See **categorical imperative**.

Imperative, hypothetical: See **hypothetical imperative**.

Individual relativism: See **ethical subjectivism**.

Infinite regress: An unending series of claims, each of which justifies a previous one and requires justification by a subsequent one. Because the chain never ends, none of the claims within it is ultimately justified.

Instrumental goods: Those things whose value consists in the fact that they help to bring about other good things. Examples include vaccinations, mothballs, and money.

Intrinsic values: Those things that are good in and of themselves, considered entirely apart from any good results they may cause. It is controversial which things are intrinsically valuable, but happiness, desire satisfaction, virtue, and knowledge are frequently mentioned candidates.

Lex talionis: The law of retaliation, the principle that says that a wrongdoer deserves to be treated just as he treated his victim.

Logical validity: The feature of an argument that indicates that its premises logically support its conclusion. Specifically, an argument is logically valid just because its conclusion must be true if its premises were all true. Another way to put this: logically valid arguments are those in which it is impossible for all premises to be true while the conclusion is false.

Maxim: A principle of action that you give to yourself. It contains your intended action and the reason you are doing it.

Metaethics: The area of ethical theory that asks about the status of normative ethical claims. It asks, for instance, about whether such claims can be true and, if so, whether personal, cultural, or divine opinion makes them true (or none of the above). It also considers issues about how to gain moral knowledge (if we can), and whether moral requirements give us reasons to obey them.

Metaphysics: The branch of philosophy that discusses the nature of reality, what exists, and what does not exist.

Monism, ethical: See **ethical monism**.

Moral agent: One who can guide his or her behavior by means of moral reasoning.

Moral community: the set of those beings whose interests are intrinsically important. Membership signifies that you are owed respect, that you have moral rights, that others owe you moral duties for your own sake.

Moral exemplar: Someone of outstanding moral character; someone who can serve as a proper moral role model.

Moral luck: A case in which the morality of an action or a decision depends on factors outside of our control.

Moral naturalism: The view that moral features are natural (i.e., not supernatural) features, whose existence can be confirmed by means of the natural sciences.

Moral nihilism: The form of moral skepticism that says that the world contains no moral features, and so there is nothing for moral claims to be true of. Its two major forms are the **error theory** and **expressivism**.

- Moral skepticism:** The view that there are no objective moral standards. Moral skepticism is also sometimes taken to refer to the view that we can have no moral knowledge.
- Moral worth:** The praiseworthy feature of an action that fulfills one's moral duty.
- Natural law theory:** The normative ethical view that says that actions are right if and only if they are natural, and wrong if and only if they are unnatural; people are good to the extent that they fulfill their true nature, bad insofar as they do not.
- Non-maleficence:** Not harming others.
- Norm:** A standard of evaluation. Norms tell us how we should or ought to behave. They represent a measure that we are to live up to.
- Normative ethics:** The area of ethical theory focused on identifying which kinds of actions are right and wrong, examining the plausibility of various moral rules, and determining which character traits qualify as virtues and which as vices.
- Normative features:** Those features that tell us how things *ought* to be, or how we *should* behave. They rely on norms to do this.
- Objective moral duties:** Those moral requirements that apply to people regardless of their opinions about such duties, and independently of whether fulfilling such duties will satisfy any of their desires.
- Objective theory of well-being:** There are many such theories, all sharing a common feature—they claim that certain things are good for people whether or not they believe them to be and whether or not such things satisfy a person's actual or informed desires.
- Occam's razor:** The instruction never to multiply entities beyond necessity. In the context of selecting from among competing theories, it tells us to choose that theory that can explain as much as any other, while making the fewest assumptions.
- Optimific:** Producing the best possible results.
- Optimific social rule:** A rule whose general acceptance within a society would yield better results than any other such rule.
- Paternalism:** The policy of treating mature people as if they were children. More specifically, it is a policy of limiting someone's liberty, against his will, for his own good.
- Pluralism, ethical:** See **ethical pluralism**.
- Premise:** Any reason that is used within an argument to support a conclusion.
- Prima facie duty:** A permanent, excellent but non-absolute reason to do (or refrain from) a certain type of action.
- Principle of utility:** The ultimate utilitarian moral standard, which says that an action is morally right if and only if it does more to improve overall well-being than any other action you could have done in the circumstances.
- Prisoner's dilemma:** A situation in which everyone involved would be better off by reducing his or her pursuit of self-interest.
- Proceduralism:** The view that says that we must follow a certain procedure in order to determine which actions are morally right, or which moral claims are true.

Psychological egoism: The view that all human actions are motivated by self-interest, and that altruism is impossible.

Punishment, exemplary: See **exemplary punishment**.

Punishment, vicarious: See **vicarious punishment**.

Relativism, cultural: See **cultural relativism**.

Relativism, individual: See **ethical subjectivism**.

Rule consequentialism: The normative ethical theory that says that actions are morally right just because they would be required by an optimific social rule.

Rule utilitarianism: The version of rule consequentialism that says that well-being is the only thing of intrinsic value.

Self-evident: A claim is self-evident just in case it is true, and adequately understanding it is enough to make you justified in believing it. The best candidates for self-evident claims are conceptual truths.

Self-regarding actions: Actions that affect only oneself.

Social contract theory: A view in political philosophy that says that governmental power is legitimate if and only if it would be accepted by free, equal, and rational people intent on selecting principles of cooperative living. Also, a view in normative ethical theory that says that actions are morally right if and only if they are permitted by rules that free, equal, and rational people would agree to live by, on the condition that others obey these rules as well.

Soundness: A special feature of some arguments. Sound arguments are ones that (1) are logically valid, and (2) contain only true premises. This guarantees the truth of their conclusions.

Standard of rightness: A rule that gives conditions that are both necessary and sufficient for determining whether actions (or other things) are morally right.

State of nature: A situation in which there is no central authority with the exclusive power to enforce its will on others.

Strictly conscientious action: Action motivated by the thought or the desire to do one's duty for its own sake, rather than from any ulterior motive.

Supererogation: Praiseworthy actions that are above and beyond the call of duty.

Tacit consent: Agreement that is expressed through silence or inaction.

Theist: One who believes that God exists.

Unconditional values: Those things that are good in every possible situation.

Universalizability: the feature of a maxim that indicates that every rational person can consistently act on it. Here is the three-part test for a maxim's universalizability: (1) carefully frame the maxim; (2) imagine a world in which everyone shares and acts on that maxim; (3) determine whether the goal within the maxim can be achieved in such a world. If so, the maxim is universalizable. If not, it isn't.

Validity: See **logical validity**.

Value theory: The area of ethics concerned with identifying what is valuable in its own right, and explaining the nature of well-being.

Veil of ignorance: An imaginary device that removes all knowledge of one's social, economic, and religious positions; one's personality traits; and other distinguishing features. It is designed to ensure that the important choices of social contractors are made fairly.

Vicarious punishment: The deliberate punishment of innocent victims, designed to deter third parties.

Vicious: Possessed of many vices. The opposite of *virtuous*.

Virtue ethics: A normative ethical theory that says that an action is morally right just because it would be done by a virtuous person acting in character.

INDEX

- abolitionism, 321. *See also* slavery
 - Bentham and, 120
 - Colfax and, 6
- abortion
 - circular reasoning and, 228
 - humanity argument and, 82–83
 - religious authority argument and, 66–67
- absolute moral rules, 174, 202–3
 - ethical monism and, 202–3
 - Kantian perspective and, 156–57, 157n1
- absolutism, 122. *See also* ethical pluralism
 - cannibalism and, 122
 - DE and, 206–11
 - irrationality and, 212–14
 - killing innocent humans and, 209, 211, 212, 213, 214, 306
 - objectivism and, 306–7
 - rape and, 306
- absolutism argument, 306–7
- abuse
 - of children, 6, 223, 321
 - Soviet psychiatric, 32–33, 37n4
- actions/intentions, utilitarianism and, 118–19, 131–32
- actual v. expected results, 117–18
- act utilitarianism. *See* utilitarianism
- Adams, John Quincy, 8
- ad hominem attack, 297
- agnostics, 59
- agony, 5
- agreement, in ethics, 14. *See also* unified moral theory
- aiming directly for happiness, 27–28
- altruism. *See also* heroism
 - carnegiehero.org and, 87
 - defined, 87
 - psychological egoism and, 86–89
 - testimony/behavior and, 98, 99
- ambiguity, 82
 - humanity and, 82, 83, 84
 - objectivity and, 276
 - sports and, 83
- Amish community, 286, 287
- amoralism, 151–52, 322n8
 - amoralists challenge and, 151–54
 - expressivism and, 301
 - Hobbe's Fool and, 190–91
- Anarchy, State, and Utopia (Nozick), 31, 37n3
- animal experimentation
 - contractarianism and, 200
 - Singer and, 120
 - utilitarianism and, 123, 124
- animals. *See also* human beings; vegetarianism
 - cheetah, 75
 - deer, 123
 - dogfighting argument and, 83
 - human beings v., 71, 72–73
 - Kantian perspective and, 172–74
 - misbehavior of, 149, 163
 - moral community and, 123–24, 200
 - moral guidance from, 11
 - natural law theory and, 71
- appearances principle, 99
- Aquinas, Thomas, 259, 273n2
- arguments, 6. *See also* validity; specific arguments
 - false beliefs and, 6, 7
 - logically valid, 7, 8, 28, 300
 - against moral objectivity, 305–22
 - sound, 9
- Aristotle. *See also* virtue ethics
 - courage and, 247
 - follow the appearances principle and, 99
 - good life and, 21
 - happiness and, 26n1
 - Nicomachean Ethics of, 241, 242, 258n3
 - Politics by, 259, 273n1

- virtue ethics and, 241
- women and, 259, 273n1
- assumptions
 - God is creator of morality, 60-65, 69
 - religion/God is source of moral guidance, 65-69
 - religious belief is needed for moral motivation, 58-60, 69
- atheism, 58
 - atheism argument, 313-14
 - error theory and, 293-94
 - objectivism and, 313-14
- athletes
 - competition and, 271
 - dogfighting argument and, 83
 - gladiators v., 116
 - good life and, 21
 - gymnast, 94, 131
 - hamstring and, 35
 - McEnroe and, 49, 55n1
 - steroids and, 179, 181, 182
- attitudinal pleasure, 19, 20. *See also* enjoyment
- autonomy, 32-34, 55
 - Kantian perspective and, 160-63, 169-71
 - rationality and, 160-63
- autonomy argument, 33-34, 169-71
- avoiding misery argument, 94-95
- babies. *See* infants
- Batman Begins*, 214
- begs the question, 82, 83, 90, 91, 92, 103
- behavior and testimony, 98, 99
- Bellow, Saul, 34
- beneficence (prima facie duty), 221
- benefit/harm ratio, 125, 128
- benevolence, 89, 98
 - Kantian perspective and, 155, 174
 - utilitarianism and, 112, 120, 130, 155, 265, 266
- Bentham, Jeremy, 120, 127. *See also* utilitarianism
- Best Argument for Ethical Egoism, 107-8
- best possible results. *See* optimific
- Bible/bibles, 66, 67, 68, 69
- Bloch, Sidney, 37n4
- brain-scanning techniques, 99n1
- Brave New World* (Huxley), 32
- breaking law. *See* civil disobedience
- breaking moral rules, 185-86, 222
- brief desires, 48-49
- The Brothers Karamazov* (Dostoevsky), 60, 313
- butterfly, happiness and, 27, 28
- cannibalism, 122
- care, 265-70
- carnegiehero.org, 87
- categorical imperatives, 153-54
- categorical reasons, 315, 316, 322n5
 - categorical reasons argument, 315-16
 - error theory and, 294, 304n1
- cheetah, 75
- children. *See also* care; infants
 - abuse of, 6, 223, 321
 - desire theory and, 46, 50
 - ethical egoism and, 104
 - golden rule and, 242
 - hedonism and, 25, 26
 - impartiality to, 133
 - Kohlberg's six-stage model and, 261
 - Leviticus and, 68
 - The Liars' Club* and, 34
 - molestation of, 223
 - moral conflict examples and, 223
 - moral education and, 245-46
 - moral responsibility and, 4
 - psychological egoism and, 96
 - Sophie's Choice* dilemma and, 250-51
 - terrorism and, 207, 209
 - tragic dilemmas argument and, 250-51
 - UNICEF and, 132, 216
 - virtue ethics and, 242, 243, 245-46, 249
- Churchill, Winston, 254. *See also* Gandhi, Mohandas
- circular reasoning, 228
- civil disobedience, 188-89, 235
- coherentism, 227-29
- Colfax, Richard, 6
- commitments, 281, 291
- common sense. *See* moral common sense
- competition, 271
 - feminist ethics and, 268-69
 - prisoner's dilemma and, 269
 - self-interest and, 178-79
- conceptual truths, 73-74
- conclusions, 6. *See also* arguments
- conflict resolution. *See* moral conflict
- conscience, guilty, 95-96
- conscientious action, strictly, 90
- conscientiousness, 59
- consent
 - contractarianism and, 194-96
 - reparation and, 105, 106
 - tacit, 194-96
- consent argument, 195-96
- consequentialism, 112-43. *See also* utilitarianism
 - difficulties in, 125-43
 - intrinsic value and, 114, 115, 123, 124
 - nature of, 114-19
 - rule consequentialism and, 140-42, 177
 - structure of, 114-15
 - utilitarianism and, 115
- conservatism, 233
- consistency. *See also* golden rule; unified moral theory
 - /fairness, Kantian perspective and, 145-47
 - golden rule and, 146
 - moral theory and, 14
 - Nazis and, 155
- continent, 247
- contractarianism (social contract theory), 176-201
 - advantages of, 183-89, 201
 - animal experimentation and, 200
 - background of, 177-78
 - civil disobedience, 188-89

- contractarianism (*continued*)
 - consent and, 194-96
 - disagreements in, 197-98
 - free-rider problem and, 191-92
 - Hobbes and, 181
 - infants and, 200
 - killing and, 176, 182, 184, 185, 200
 - law and, 186-89
 - moral community and, 198-200
 - morality in, 183
 - moral rules and, 183-86
 - political theory and, 177-78
 - prisoner's dilemma and, 178-83
 - problems with, 190-201
 - proceduralism and, 176-77, 184
 - punishment and, 186-87
 - Rawls and, 183-84, 197, 268
 - self-interest and, 198-199
 - Socrates and, 178
 - status of morality and, 185
 - why be moral?* and, 190-94
- contradiction
 - cultural relativism and, 286-88, 291
 - ethical pluralism and, 211-12, 219
 - ethical subjectivism and, 283-84, 291
 - expressivism and, 299
 - virtue ethics and, 254-56
- contradiction* argument, 212, 219
- conventional moral wisdom. *See* moral common sense
- cooperation
 - feminist ethics and, 271-72
 - state of nature and, 181-83
- core moral beliefs. *See* moral common sense
- courage, 164, 247
- cultural relativism, 278-91
 - commitments and, 281, 291
 - contradiction and, 286-88, 291
 - disagreement and, 285-86, 291
 - ideal observers and, 288-91
 - intrinsic value and, 280-81
 - moral equivalence and, 280, 291
 - moral infallibility and, 279-80, 291
 - moral progress and, 281-83, 291
- DDA. *See* doctrine of doing and allowing
- DDE. *See* doctrine of double effect
- death penalty, 113, 167
- decision procedure
 - for prima facie duties, 231-33
 - rightness standard v., 130-31
 - utilitarianism and, 232
- deer, hiker and, 123
- definitions, natural law theory and, 84-85
- deists, 61, 66
- deliberation, utilitarianism and, 129
- demanding nature
 - of utilitarianism, 129-32
 - of virtue ethics, 253
- dependence on others, 263-64
- desires
 - brief, 48-49
 - deepest, fallibility of, 52-53
 - false beliefs and, 46-47
 - Hume and, 165
 - impoverished, 50-51
 - informed, 47, 50, 52, 54
 - mild, 48-49
 - satisfaction, ignorance of, 50
 - strongest desires* argument, 89-92
- desire satisfaction theory, 38-55
 - attractions of, 39-44
 - central claims of, 45
 - children and, 46, 50
 - definition of, 38-39
 - and knowledge of good, 43-44
 - mentally retarded and, 46, 53, 54
 - motivation and, 41-42
 - objectivism and, 40-41, 43, 44, 54, 55
 - personal authority and, 39-40
 - problems for, 45-55
 - self-harm and, 51
 - self-interest and, 42-43
 - self-sacrifice and, 51
 - slavery and, 51, 54
 - suicide and, 46, 52
- dilemma, prisoner's. *See* prisoner's dilemmas
- direct pursuit of happiness, 27-28
- disagreement
 - cultural relativism and, 285-86, 291
 - disagreement* argument, 312-13
 - ethical subjectivism and, 284-85, 291
 - expressivism and, 299
- disappointment, 49, 91
- disaster prevention* argument, 205-6, 208, 213, 219.
 - See also irrationality* argument
- DDE and, 208
- ethical pluralism and, 205-6, 208, 213
- prima facie duties and, 224-25
- disastrous results* argument, 295-96
- discipline, ethics and, 3
- disease, 30
- divine command theory, 61-65, 70n3, 290
 - defined, 61
 - Euthyphro* argument and, 63-64, 290, 291n5
 - Socrates and, 61-62, 291
 - virtue ethics v., 257
- divine perfection* argument, 64
- "Do all the good you can..." motto, 112, 131
- Doctorow, E. L., 136
- doctrine of doing and allowing (DDA), 214-18
- doctrine of double effect (DDE), 206-11
- dogfighting* argument, 83
- dogmatism, 309-10
- dogmatism* argument, 309-10
- doing and allowing doctrine. *See* doctrine of doing and allowing
- Donner party, 122
- Dostoevsky, 60, 65, 313
- double effect doctrine. *See* doctrine of double effect
- The Drowned and the Saved* (Levi), 81, 85n2
- drugs
 - addiction, 91, 243

- heroin* argument, 7
 performance-enhancing, 179, 181, 182
 religion as, 294
 as scarce resource, 179
 Schwartz and, 34
 scientific testing of, 217
 soma pill, 32, 35
 Soviet psychiatric abuse and, 32-33, 37n4
- education. *See* moral education
 Efficiency Model, 79, 80, 81
 egoism. *See* ethical egoism; psychological egoism
 elves hypothesis, 97-98
 emergency room example, 206-7
Emile: On Education (Rousseau), 259, 273n4
 emotions
 care and, 266-67
 feminist ethics and, 271
 Kantian perspective and, 165, 266, 273n6
 moral understanding and, 244
 /reasons, moral knowledge and, 164-65
 empirical truths, 73, 74
 enjoyment. *See also* happiness
 attitudinal pleasure as, 19, 20
 happiness as, 20, 26n1
 Epicurus, 21. *See also* hedonism
equal rights argument, 308-9
 equivalence. *See* moral equivalence
 error theory, 292-97, 322n6
 atheism and, 293-94
 categorical reasons and, 294, 304n1
 disastrous results argument and, 295-96
 expressivism v., 297-98, 299
 implications of, 295-96
 as metaphysical theory, 296
 moral theory v., 296
 rape and, 297
 untrustworthiness and, 296-97
 eternal damnation, 102
 ethical egoism, 100-11. *See also* psychological egoism
 children and, 104
 consent/reparation and, 105, 106
 ethical egoism argument and, 107-8
 immorality and, 101, 102, 103, 104
 killing and, 101, 102, 103, 104, 109, 111
 libertarian argument and, 105-6
 moral common sense and, 109
 moral rights and, 109-10
 paradigm cases argument and, 101-2, 103
 problems for, 108-11
 psychological egoism v., 100
 self-interest and, 101, 104, 110-11
 self-reliance argument and, 104-5
 why be moral? and, 101-4, 106-8
 ethical monism, 202-3. *See also* utilitarianism
 absolute moral rules and, 202-3
 assumptions in, 4-5
 attractions of, 203, 219n1
 ethical objectivism. *See* objectivism
 ethical particularism, 233-39
 ethical pluralism, 202-39. *See also* ethical particularism; virtue ethics
 contradiction and, 211-12, 219
 DDA and, 214-18
 DDE and, 206-11
 defined, 203
 disaster prevention argument and, 205-6, 208, 213
 ethical particularism and, 233-38
 moral conflict and, 211-12
 prima facie duties and, 220-33
 terrorists and, 203-6
 torture and, 203-6
 ethical relativism, 276-91. *See also* cultural relativism; ethical subjectivism
 defined, 277
 kinds of, 278-79
 moral nihilism v., 292
 moral skepticism and, 276-78, 291n1
 objectivism and, 292
 ethical subjectivism (individual relativism), 278-91
 commitments and, 281, 291
 contradiction and, 283-84, 291
 disagreement and, 284-85, 291
 ideal observers and, 288-91
 intrinsic value and, 280-81
 moral equivalence and, 280, 291
 moral infallibility and, 279-80, 291
 moral progress and, 281-83, 291
 ethical theories. *See* moral theories
 ethics (moral philosophy). *See also* feminist ethics;
 good life; morality; virtue ethics
 agreement in, 14
 care and, 265-70
 complexity of, 3
 core areas of, 1-3
 discipline and, 3
 introduction, 1-15
 metaethics and, 2, 11
 moral philosophy as, 1
 normative, 2
 starting points, 3-5
 value theory and, 1, 2
 vastness of, 1
 eudaimonia, 248-49. *See also* happiness
Euthyphro (Plato), 61-62
Euthyphro argument, 63-64, 290, 291n5. *See also*
 divine command theory
 evaluative beliefs, 318
evil pleasures argument, 28-29, 31, 37
 exemplary punishment, 136-37
 expanding realm of self-interest (strategy),
 96-97
expected benefit argument, 92-94
expected v. actual results, 117-18
 "experience machine," 31
 experiences of women, feminist ethics and, 260,
 262-65
 expressivism, 292, 297-304, 322n7
 eye-for-an-eye principle. *See* lex talionis

- factual claims, moral claims v., 298–299
 failure of logic, 7
 fairness. *See also* universalizability principle
 /consistency, Kantian perspective and, 145–47
 fallibility of deepest desires, 52–53
 false beliefs
 arguments and, 6, 7
 desires and, 46–47
 happiness and, 31–32, 37
 suicide and, 52
 false happiness, 31–32, 37
 fanatics, 146–47, 155, 161
 fancies, passing, 48–49. *See also* desires
Fathers and Sons (Waugh), 143n2
 fear of God, 59, 60
 feminist ethics, 259–73
 care and, 265–70
 central claims of, 260
 challenges for, 270–72
 competition and, 268–69
 cooperation and, 271–72
 elements of, 259–60
 emotions and, 271
 impartiality and, 268, 271
 justice and, 269–70, 272
 lex talionis and, 268
 moral community and, 270–71
 moral development and, 261–62
 moral reasoning and, 260, 261–62, 277
 moral rights and, 269–70, 272
 Rawls and, 268
 unified moral theory and, 267–68, 270, 271
 women's experiences and, 260, 262–65
 fidelity (prima facie duty), 221
 Fireside, Harvey, 37n4
 first motivational argument, 41–42
 fitness, 80
 Fitness Model, 79, 80, 81
 fMRI, 99n1
 follow the appearances principle, 99
 Fool (Hobbes), 190–91
 Foot, Philippa, 25, 26n2
 foresight/intention, DDE and, 209–11
 Fraser, Flora, 175
 free-rider problem, 191–92
 Freud
 Kantian perspective and, 165
 religion and, 294
 sexual impulses and, 13
 fulfillment of desires. *See* desire satisfaction theory
 fundamental moral rules, knowledge of, 226–30

 Gandhi, Mohandas, 188, 247, 253, 254
 Gates, Bill, 178
 gays. *See* homosexuals
 generality, 233
 getting what you want. *See* desire satisfaction theory
gild and stain, 277
 Gilligan, Carol, 261–62, 272, 273n5
 gladiators, 83, 116, 119
 God, 70n2. *See also* atheism; divine command theory; religion
 deists and, 61, 66
 eternal damnation and, 102
 fear of, 59, 60
 God is creator of morality (assumption), 60–65, 69
 God/religion is source of moral guidance (assumption), 65–69
 “If God is dead, then everything is permitted,” 60, 313
 intelligent designer and, 78, 79
 justice and, 59
 lex talionis and, 68
 perfection of, 64–65
 rape and, 62, 64, 69, 70
 religion/morality and, 58–70
 religious belief is needed for moral motivation (assumption), 58–60, 69
 torture and, 63, 64, 69
 golden rule, 2, 145, 146–47
 categorical imperatives and, 153
 children and, 242
 definition of, 146
 proceduralism and, 177
 what if everyone did that? test and, 145, 146, 149, 155, 232
 good. *See also* greatest good
 “Do all the good you can...” motto, 112, 131
 knowledge of, 43–44
 maximizing goodness principle, 115–16
 good life. *See also* desire satisfaction theory;
 hedonism
 athletes and, 21
 desire satisfaction theory and, 38–55
 exceptions to rules of, 24
 happiness and, 19, 22–23, 31–32
 hedonism and, 18–37
 instrumental goods and, 18, 20
 intrinsic value of, 19, 20, 23
 misery and, 22–23
 models of, 21–22, 39
 objective values and, 55
 philosophy and, 21, 22
 Plato and, 21
 rules of, 24
 standard for, 18
 value theory and, 1, 2
 virtue and, 248–49
 good will, 163–65, 318
 gratitude (prima facie duty), 221
 Grayson, William John, 6, 7
 greatest good (*summum bonum*)
 for greatest number, 115, 130
 happiness and, 22, 28
 philosophy and, 22
 guidance. *See* moral guidance
 guilty conscience, 95–96
 Gyges, 86
 gymnast, 94, 131

- hamstring, marathon runner and, 35
- happiness. *See also* unhappiness; well-being
 aiming directly for, 27–28
 Aristotle and, 26n1
 butterfly and, 27, 28
 as enjoyment, 20, 26n1
 eudaimonia as, 248–49
 false, 31–32, 37
 good life and, 19, 22–23, 31–32
 greatest good and, 22, 28
 intrinsic value of, 18–20
 lobotomy and, 25, 26, 33, 54
 loved ones and, 25–26
 unconditional value and, 20, 163
- Hare, Richard, 158–59, 175n1
- harm
 harm/benefit ratio, 125, 128
 harms argument and, 36
 preventing (prima facie duty), 221, 225
 unhappiness and, 35–36
- hedonism, 18–37. *See also* Mill, John Stuart
 attractions of, 21–26
 autonomy and, 32–34
 children and, 25, 26
 criticism of, 27–37
 defined, 19
 Epicurus and, 20
 evil pleasures argument and, 28–29, 31, 37
 false happiness and, 31–32, 37
 Foot and, 25, 26n2
 hedonist's test and, 26
 homosexuals and, 26
 Mill and, 21, 115–16
 moral common sense and, 24
 paradox of, 27–28, 37, 130, 143n1
 physical/sensual pleasure and, 19, 20, 21
 popularity of, 36–37
 Ross and, 29–30
 two worlds argument and, 29–30, 37
 utilitarianism and, 115–16
- heroin argument, 7
- heroism. *See also* altruism
Batman Begins and, 214
 carnegiehero.org and, 87
 ethical egoism and, 107
One Flew over the Cuckoo's Nest and, 33
 revolutionary, 172
- hiker, 123, 316
- Hitler, 78, 242, 280
- Hobbes, Thomas. *See also* contractarianism
 contractarianism and, 181
 Fool of, 190–91
 human nature and, 76
 rationality and, 190–94, 199
 why be moral? and, 190–94
- homelessness, 91, 116
- homosexuals
 hedonism and, 26
 marriage argument and, 84
 morality and, 145, 146
 natural law theory and, 80, 81, 84
- Hooker, Brad, 141
- Hornby, Nick, 130
- How to Be Good* (Hornby), 130
- human beings. *See also* innocent human beings
 ambiguity of term, 82, 83, 84
 animals v., 71, 72–73
 as biological category, 82
 as moral category, 82–83
 natural law theory and, 72–73
- humanity argument, 82–83
- humanity principle (Kant)
 description of, 159–61
 just deserts and, 166–69
 moral community and, 172–74
 problems with, 165–74
 vagueness and, 166
- human nature, 75
 conceptions of, 75–78
 infants and, 76
 innate traits and, 75–76
 mentally retarded and, 77
 as objective standard of morality, 72
 scientific study of, 74–75
 universal human traits and, 77–78
- human traits, 75–78
- Humboldt's Gift* (Bellow), 34
- Hume, David, 73, 74
 desires and, 165
 Hume's argument, 73–74, 76
 moral nihilism and, 277
 motivational argument and, 316–18
Treatise on Human Nature, 85n1
- hunger strikes, 253, 254
- hunting example, 110
- Huxley, Aldous, 32
- hypothetical imperatives, 152
- iconoclast, 279
- ideal observers, 288–91
- "If God is dead, then everything is permitted,"
 60, 313
- ignorance
 of desire satisfaction, 50
 veil of, 184
- immorality. *See also* ethical egoism
 ethical egoism and, 101, 102, 103, 104
 fanatics and, 161
 irrationality and, 151
 irrationality of immorality argument, 153–54
 Plato and, 102
 rape and, 76, 81, 104, 120, 134, 161
 rewards of, 101
 slavery and, 161
 torture and, 203–5
 torture is immoral (example) and, 298–299, 300, 301, 302
- impartiality
 feminist ethics and, 268, 271
 utilitarianism and, 119–20, 133–34, 143n2
- imperatives, 152–54
- implications of egoism argument, 88–89

I-7 INDEX

- impoverished desires, 50-51
- In a Different Voice* (Gilligan), 261-62, 272, 273n5
- individual relativism. *See* ethical subjectivism
- infallibility, moral, 279-80, 291
- infants. *See also* children
 - actions of, 149, 172
 - care and, 266
 - contractarianism and, 200
 - gentleness trait and, 74-75
 - human nature and, 76
 - Kantian perspective and, 172-74
 - moral community and, 124, 173
- infinite regress, 227, 231
- informed desires, 47, 50, 52, 54
- informed requests, 5
- injustice
 - correctness of, 138, 143n3
 - optimific and, 138-39
 - utilitarianism and, 136-39
- innate traits, 75-76
- innocent human beings. *See also* abortion
 - betrayal of, 114
 - Kantian perspective and, 156, 157
 - killing, absolutism and, 209, 211, 212, 213, 214, 306
 - killing, prima facie duties and, 237
 - killing, *tragic dilemmas* argument and, 250-51
 - killing, utilitarianism and, 118, 120, 122, 126, 135-36, 138, 139, 142
 - proceduralism and, 176
 - saving, ethical pluralism and, 204, 205, 206, 207, 208
 - saving, priority problem and, 257
 - switchman example and, 217
- instrumental goods, 18, 20
- integrity, 155. *See also* Kantian perspective
- intelligent designer, 78, 79. *See also* God
- intentional killing, 74
- intentions/actions, utilitarianism and, 118-19, 131-32
- intentions/foresight, DDE and, 209-11
- interests of others, self-interest v., 42
- interpretation of Bible/bibles, 66, 67, 68, 69
- intolerance, 280, 310-12
- intrinsically bad, 19, 114, 115
- intrinsic values, 19
 - cultural relativism and, 280-81
 - ethical subjectivism and, 280-81
 - good life and, 19
 - of happiness, 18-20, 23
 - utilitarianism and, 114, 115, 123, 124, 134-35
- invisible elves hypothesis, 97-98
- irrationality
 - absolutism and, 212-14
 - immorality and, 151
 - of terrorists, 151
- irrationality* argument, 213-14, 224-25
- irrationality of immorality* argument, 153-54
- is v. ought*, 74, 75, 78, 85. *See also* natural law theory
- Johnson, Samuel, 198
- just deserts, humanity principle and, 166-69
- justice, 4, 5. *See also* Kantian perspective
 - death penalty and, 113
 - feminist ethics and, 269-70, 272
 - God and, 59
 - Kantian perspective and, 143, 174
 - as prima facie duty, 221
 - taxation and, 122
 - Theory of Justice* and, 183-84
 - utilitarianism and, 136-39
- justifying moral rules, 184-85
- Kant, Immanuel, 144
 - Observations on the Feeling of the Beautiful and the Sublime* by, 259, 273n3
 - women and, 259, 273n3
- Kantian perspective, 144-75
 - absolute moral rules and, 156-57, 157n1
 - autonomy and, 160-63, 169-71
 - benevolence and, 155, 174
 - emotions and, 165, 266, 273n6
 - fairness/consistency and, 145-47
 - Freud and, 165
 - humanity principle and, 159-61, 165-74
 - infants and, 172-74
 - innocent human beings and, 156, 157
 - integrity and, 155, 174
 - justice and, 143, 174
 - killing and, 161, 162, 163, 168, 171
 - mentally retarded and, 172-74
 - moral community and, 172-74
 - morality/rationality and, 151-54
 - slavery and, 158-59, 161
 - taxation example and, 144, 147
 - universalizability principle and, 147-50, 232, 271
 - utilitarianism v., 163, 174-75
- Kaplan, James, 55n1
- Karamazov, Ivan, 60, 313
- karma doctrine, 102
- Karr, Mary, 34
- Kesey, Ken, 33
- Khalid Sheikh Mohammed. *See* Mohammed, Khalid Sheikh
- killing. *See also* abortion; death penalty; innocent human beings; *meat eating* argument; rape; suicides; torture; vegetarianism
 - contractarianism and, 176, 182, 184, 185, 200
 - DDA and, 214, 215, 216, 217, 218
 - ethical egoism and, 101, 102, 103, 104, 109, 111
 - Gyges and, 86
 - intentional, 74
 - Kantian perspective and, 161, 162, 163, 168, 171
 - religion and, 59, 62, 64, 68
 - skin color and, 76
 - V2 rocket and, 78
- kindness. *See* benevolence
- King, Martin Luther, Jr., 188
- Kohlberg, Lawrence, 261
- Kohlberg's six-stage model, 261

- latrine cleaner, 53
- law. *See also* lex talionis
 civil disobedience and, 188-89, 235
 contractarianism and, 186-89
 obeying, 186
- lesbians. *See* homosexuals
- Levi, Primo, 81, 85n2
- Leviathan* (Hobbes), 181, 191-92
- Leviticus, 68
- lex talionis (eye-for-an-eye principle), 68, 167-69, 268
- The Liars' Club* (Karr), 34
- libertarian argument, 105-6
- libertarians, 105, 106
- life's trajectory* argument, 34-35
- lobotomy, 25, 26, 33, 54
- logic, 7, 9
- logically valid arguments, 7, 8, 28, 300
- Lord of the Flies*, 181
- loved ones, happiness and, 25-26
- luck, moral, 166, 171-73
- lying
 allowance for, 123, 306
 to investigator, 101
 Kantian perspective and, 150, 156, 157
 to ourselves, 301, 303
 totalitarian regimes and, 135
- lying promise, 150
- marathon runner, hamstring and, 35
- The March* (Doctorow), 136
- marriage argument, 84
- Marx, religion and, 294
- maximizing goodness* principle, 115-16
- maxims, 147-50. *See also* universalizability principle
- McEnroe, John, 49, 55n1
- McMurphy, R. P., 33
- meat eating* argument, 9-11
- mentally retarded
 contractarianism and, 200
 desire theory and, 46, 53, 54
 human nature and, 77
 Kantian perspective and, 172-74
 moral community and, 172, 173
 utilitarianism and, 124
- metaethics, 2, 11. *See also* morality, status of
- metaphysical theories, 296. *See also* error theory
- might v. right, 5
- mild desires, 48-49
- Mill, John Stuart
 Bentham and, 120, 127
 deliberation and, 129
 hedonism and, 21, 115-16
maximizing goodness principle and, 115-16
 pleasures and, 127
 Socrates and, 127
On the Subjection of Women by, 120
 taxation and, 121-22
 utilitarianism and, 115, 121-22, 127, 129
Utilitarianism by, 121-22
- misbehavior, of animals, 149, 163
- misery
 autonomy and, 32
avoiding misery argument, 94-95
 disappointment and, 49, 91
 disease and, 30
 good life and, 22-23
 hedonism and, 20, 22-23
trajectory argument and, 34-35
- Mohammed, Khalid Sheikh, 203-4
- molester, child, 223
- monism. *See* ethical monism
- Moore, G. E., 112, 113, 124n1
- moral absolutism. *See* absolutism
- moral beliefs. *See* moral common sense
- moral claims
 conceptual truths and, 73-74
 empirical truths and, 73, 74
 expressivism and, 301-3
 factual claims v., 298-299
- moral common sense
 ethical egoism and, 109
 hedonism and, 24
 utilitarianism and, 120-21
- moral community
 animals and, 123-24, 200
 contractarianism and, 198-200
 defined, 123
 feminist ethics and, 270-71
 humanity principle and, 172-74
 infants and, 124, 173
 Kantian perspective and, 172-74
 mentally retarded and, 172, 173
 utilitarianism and, 123-24, 143, 175
- moral complexity, virtue ethics and, 242-43
- moral conflict
 ethical pluralism and, 211-12
 prima facie duties and, 223-24
 resolution, utilitarianism and, 121-22
 virtue ethics and, 254-56
- moral development, feminist ethics and, 261-62
- moral disagreement. *See* disagreement
- moral education, 245-46
- moral equivalence
 cultural relativism and, 280, 291
 ethical subjectivism and, 280, 291
 expressivism and, 299-300
- moral exemplar, 241
- moral flexibility, utilitarianism and, 122-23
- moral guidance
 from animals, 11
 religion/God is source of moral guidance (assumption), 65-69
 universal human traits and, 77-78
 virtue ethics and, 251-53
- moral importance, permanent, 237-38
- moral infallibility, 279-80, 291
- morality. *See also* immorality; *why be moral?*
 contractarianism and, 183
 creation of, 61
 norms and, 60

- morality (*continued*)
 origins of, natural law theory and, 73
 /rationality, Kantian perspective and, 151-54
 religion and, 58-70
 as social phenomenon, 183, 201
 morality, status of (metaethics), 274-322. *See also*
 objectivism
 contractarianism and, 185
 ethical relativism and, 276-91
 moral nihilism and, 292-304
 objectivism and, 305
 moral knowledge
 ethical particularism and, 236-37
 natural law theory and, 73-75
 reason v. emotions in, 164-65
 utilitarianism and, 116-17
 moral luck, 166, 171-72
 moral motivation. *See also motivational arguments*
 desire satisfaction theory and, 41-42
 religious belief is needed for moral motivation
 (assumption), 58-60, 69
 utilitarianism and, 129-31
 moral naturalism, 320-21
 moral nihilism, 292-304. *See also* error theory;
 expressivism
 defined, 277, 292
 ethical relativism v., 292
 objectivism and, 292
 moral objectivism. *See* objectivism
 moral philosophy. *See* ethics
 moral progress, 281-83, 291
 moral reasoning, 5-12
 /emotions, moral knowledge and, 164-65
 feminine ethics and, 260, 261-62, 277
 goals of, 6
 tests for, 6-7
 women and, 260, 261-62, 272
 moral regret test, 224, 237
 moral rights
 equal rights argument, 308-9
 ethical egoism and, 109-10
 feminist ethics and, 269-70, 272
 moral rules. *See also* absolute moral rules
 breaking, 185-86, 222
 content of, 183-84
 contractarianism and, 183-86
 fundamental, knowledge of, 226-30
 justifying, 184-85
 self-regarding, 183
 moral skepticism
 ethical relativism and, 276-78, 291n1
 fundamental moral rules and, 227
 moral theories, 12-14. *See also* unified moral theory
 consistency and, 14
 difficulties in, 27
 error theory v., 296
 testing, self-evidence and, 230-31
 as work in progress, 201
 moral understanding, 243-44. *See also* moral
 common sense
 moral worth, 164-65. *See also* good will
 mortality argument, 300
 mother's care. *See* care
 motivation. *See* moral motivation
 motivational arguments, 41-42
 Hume and, 316-18
 objectivism and, 316-18
 MRI, functional, 99n1
 multiple harms argument, 36
 murder. *See* killing
 musicians, 21, 40, 49, 53, 54

 Natural Goodness (Foot), 25, 26n2
 naturalism, moral, 320-21
 natural law argument, 81-82
 natural laws, 85
 natural law theory, 71-85
 animals and, 71
 definitions and, 84-85
 is v. ought, 74, 75, 78, 85
 morality origins and, 73
 moral knowledge and, 73-75
 objective moral standards and, 72
 suicide and, 80, 81
 natural purposes, 78-82
 nature, state of, 181-83
 navigator of ship (example), 245-46
 Nazis
 consistency and, 155
 DDA and, 215-16
 as fanatics, 146, 147, 155
 Levi and, 81, 85n2
 moral progress after, 282
 weapons intelligence and, 137
 Nicomachean Ethics (Aristotle), 241, 242, 258n3
 nihilism. *See* moral nihilism
 Nolan, Christopher, 214
 non-maleficence (*prima facie* duty), 221
 normative ethics, 2
 normative features, 319
 norms, 60, 319
 Nozick, Robert, 31, 37n3

 objective moral standards, 72, 276
 objective truths argument, 307
 objective values
 avoiding, 40-41
 good life and, 55
 science and, 318-21
 objectivism (moral/ethical objectivism), 276,
 305-22. *See also* cultural relativism; ethical
 subjectivism
 absolutism and, 306-7
 arguments against, 305-22
 atheism and, 313-14
 categorical reasons argument and, 315-16
 defined, 277
 desire satisfaction theory and, 40-41, 43, 44,
 54, 55
 disagreement argument and, 312-13
 dogmatism argument and, 309-10
 equal rights argument and, 308-9

- ethical relativism and, 292
- intolerance and, 310-12
- moral nihilism and, 292
- motivational argument and, 316-18
- objective truths argument and, 307
- scientific test of reality argument and, 318-21
- status of ethics and, 305
- tolerance argument and, 310-12
- objectivity, 276
- Observations on the Feeling of the Beautiful and the Sublime* (Kant), 259, 273n3
- Occam's razor, 318, 320, 321
- One Flew over the Cuckoo's Nest* (Kesey), 33
- On the Subjection of Women* (Mill), 120
- opiate, religion as, 294
- optimific, 113, 114, 117, 118, 119, 125, 128, 129, 131, 133, 134, 136, 142. *See also* utilitarianism
- injustice and, 138-39
- social rules, 140-42, 149
- ought v. is, 74, 75, 78, 85. *See also* natural law theory
- paradigm, 102
- paradigm cases argument, 101-2, 103
- paradox of hedonism, 27-28, 37, 130, 143n1
- paradox of self-harm, 51
- paradox of self-sacrifice, 51
- parsimony, 233
- partiality, 133. *See also* impartiality
- particularism. *See* ethical particularism
- passing fancies, 48-49. *See also* desires
- paternalism, 33, 161, 162
- Pauline Bonaparte: Venus of Empire* (Fraser), 175n2
- perennial questions
 - knowledge of good, 43-44
 - why be moral?, 101-4, 106-8, 190-94
- perfection, of God, 64-65
- performance-enhancing drugs, 179, 181, 182
- permanent moral importance, particularism and, 237-38
- personal authority
 - desire satisfaction theory and, 39-40
 - well-being and, 22
- philosophy, 21, 22. *See also* ethics
- Philosophy and Public Affairs* 8 (1979), 175n1
- physical pleasure, 19, 20, 21
- piety, 61, 62
- Plato. *See also* Socrates
 - Euthyphro* by, 61-62
 - good life and, 21
 - immorality and, 102
 - Republic* by, 86, 102, 178
 - virtue and, 248
- pleasures. *See also* enjoyment
 - attitudinal, 19, 20
 - evil, 28-29, 31, 37
 - Mill and, 127
 - physical/sensual, 19, 20, 21
- plumber example, 159, 160, 166
- political theory, contractarianism and, 177-78
- Politics* (Aristotle), 259, 273n1
- premises, 6. *See also* arguments
- preventing harm (prima facie duty), 221, 225
- prima facie duties, 220-33, 238
 - advantages of, 222-25
 - decision procedure for, 231-33
 - disaster prevention argument and, 224-25
 - ethical particularism v., 233-34
 - irrationality argument and, 224-25
 - killing innocent humans and, 237
 - moral conflict and, 223-24
 - problems for, 225-26, 238
 - promise-keeping and, 230, 232, 234, 237
 - ranking of, 225-26
 - self-evidence of, 229-30
 - types of, 221
 - utilitarianism and, 224-25
- Principia Ethica* (Moore), 112, 113, 124n1
- principle of humanity. *See* humanity principle
- principle of utility. *See* utility principle
- Principles of Morals and Legislation* (Bentham), 120
- priority problem, virtue ethics and, 256-58
- prisoner's dilemmas, 178-81
 - competition and, 269
 - state of nature as, 181-83
- proceduralism, 175-77, 184. *See also* contractarianism; golden rule; rule consequentialism; universalizability principle
- promises
 - lying promise and, 150
 - promise-keeping and, 230, 232, 234, 237
- psychiatric abuse, Soviet, 32-33, 37n4
- psychological egoism. *See also* ethical egoism
 - altruism and, 86-89
 - avoiding misery argument and, 94-95
 - children and, 96
 - defined, 86-87
 - ethical egoism v., 100
 - expanding realm of self-interest and, 96-97
 - expected benefit argument and, 92-94
 - falseness of, 98-99
 - guilty conscience and, 95-96
 - implications of egoism argument, 88-89
 - strongest desires argument and, 89-92
 - testimony/behavior and, 98, 99
- psychological egoism argument, 88-89
- pulling plug, ventilator and, 218
- punishment. *See also* lex talionis
 - contractarianism and, 186-87
 - exemplary, 136-37
 - vicarious, 136-37
- rape
 - absolutism and, 306
 - correctness of, 206
 - DDA and, 215, 218
 - deliberation and, 129
 - error theory and, 297
 - God and, 62, 64, 69, 70
 - immorality and, 76, 81, 104, 120, 134, 161
 - irrationality and, 151

- rape (*continued*)
 lex lex talionis and, 168, 169
 "regret test" and, 237
 self-interest and, 102, 104, 109, 176
 virtue ethics and, 256, 257
 vulnerability, women, 262, 263
- rationality
 /autonomy, Kantian perspective and, 160-63
 Hobbes and, 190-94, 199
 /morality, Kantian perspective and, 151-54
 Rawls, John, 183-84, 197, 268. *See also*
 contractarianism
 feminist ethics and, 268
Theory of Justice by, 183-84
- razor, Occam's, 318, 320, 321
- reasoning, circular, 228. *See also* moral reasoning
- reasons. *See* premises
- Reddaway, Peter, 37n4
- Red Sox victory, 47, 48
- regress, infinite, 227, 231
- "regret test," 224, 237
- relativism. *See* ethical relativism
- religion. *See also* atheism; God
 Bible/bibles, 66, 67, 68, 69
 Freud and, 294
 God/religion is source of moral guidance
 (assumption), 65-69
 Marx and, 294
 and morality, 58-70
 as opiate, 294
 religious authority argument and, 66-67
 religious belief is needed for moral motivation
 (assumption), 58-60, 69
- reparations
 consent and, 105, 106
 as prima facie duty, 221
- Republic* (Plato), 86, 102, 178
- retaliation law. *See* lex talionis
- retarded. *See* mentally retarded
- revolutionary hero, 172
- right action standard, 241-42
- rightness standard, decision procedure
 v., 130-31
- rights. *See* moral rights
- right v. might, 5
- role models, for virtue ethics, 254, 256
- Ross, W. D., 220. *See also* ethical pluralism; hedonism
 hedonism and, 29-30
 prima facie duties and, 220-33
- Rousseau, Jean-Jacques
Emile: On Education by, 259, 273n4
 human nature and, 76
 women and, 259, 273n4
- rule consequentialism, 140-42, 177
- rules. *See* moral rules
- sadist, 20, 114
- sadness. *See* misery
- saints, 99, 129, 130, 131
- same-sex marriage, 84. *See also* homosexuals
- satisfied desires. *See* desire satisfaction theory
- Schwartz, Delmore, 34
- science
 drugs and, 217
 human nature and, 74-75
 objective values and, 318-21
 theories and, 231
- scientific test of reality* argument, 318-21
- Second Amendment, 67, 68
- second motivational* argument, 41-42
- Seinfeld*, 162
- self-evidence
 of prima facie duties, 229-30
 testing moral theories and, 230-31
- self-harm, 51
- self-improvement (prima facie duty), 221
- self-interest, 5. *See also* altruism; ethical egoism;
 psychological egoism
 competition and, 178-79
 contractarianism and, 198-199
 desire satisfaction theory and, 42-43
 ethical egoism and, 101, 104, 110-11
 expanding realm of, 96-97
 interests of others v., 42
 prisoner's dilemma and, 178-83
 rape and, 102, 104, 109, 176
- self-regarding actions, 147
- self-regarding desires, 48, 49, 50
- self-regarding moral rules, 183
- self-reliance* argument, 104-5
- self-sacrifice, 51
- sensual pleasure, 19, 20, 21
- September 11 attacks, 203-4
- Sheikh Mohammed. *See* Mohammed, Khalid
- Sheikh
 shepherd (Gyges), 86
 Sherman's command, 136, 141
- ship navigator example, 245-46
- The Simpsons*, 248
- Singer, Peter, 120
- skepticism. *See* moral skepticism
- slavery
 desire satisfaction theory and, 51, 54
 Hare and, 158-59, 175n1
 immorality and, 161
 Kantian perspective and, 158-59, 161
 moral naturalism and, 321
 slavery argument and, 6-7
- Smart, J. J. C., 142
- social contract theory. *See* contractarianism
- social phenomenon, morality as, 183, 201
- Socrates. *See also* Plato
 contractarianism and, 178
 divine command theory and, 61-62, 291
- Euthyphro* and, 61-62
- Mill and, 127
- mortality argument and, 300
- soma pill, 32, 35
- Sophie's Choice* dilemma, 250-51
- The Sopranos*, 201n1
- sound arguments, 9
- Soviet psychiatric abuse, 32-33, 37n4

- Soviet Psychiatric Abuse: The Shadow over World Psychiatry* (Reddaway & Bloch), 37n4
- Soviet Psychoprisons* (Fireside), 37n4
- sports. *See also* athletes
- ambiguity of term, 83
 - dogfighting* argument and, 83
 - gladiators and, 83
 - McEnroe and, 49, 55n1
 - women and, 269
- stain, gild* and, 277
- Stalin
- ethical subjectivism and, 280
 - evil nature of, 320
 - idolization of, 242
 - stranger as, 117
- standard of rightness, decision procedure *v.*, 130-31
- starting points, ethical, 3-5
- state of nature, 181-83
- status of morality. *See* morality, status of
- steroids, athletes and, 179, 181, 182
- strictly conscientious action, 90
- strongest desires* argument, 89-92
- subjectivism. *See* ethical subjectivism
- subjectivity of truth, 307
- suicides
- desire theory and, 46, 52
 - false belief and, 52
 - natural law theory and, 80, 81
 - self-regarding actions and, 147
 - suicide bombers and, 254
- Summa Theologica* (Aquinas), 259, 273n2
- summum bonum*. *See* greatest good
- supererogation, 132
- switchman example, 217
- tacit consent, 194-96
- taxation
- Kantian perspective and, 144, 150
 - Mill and, 121-22
- terrorism/terrorists
- children and, 207, 209
 - courage and, 164
 - DDA and, 215, 218
 - ethical pluralism and, 203-6
 - integrity and, 155
 - irrationality of, 151
 - punishment of, 136, 138, 207, 209
 - torture of, 204-6
- testability, 233
- testimony and behavior, 98, 99
- tests. *See specific tests*
- theism
- defined, 62
 - as metaphysical theory, 296
- theories. *See also* divine command theory; error theory; moral theories
- metaphysical, 296
 - scientific, 231
- Theory of Justice* (Rawls), 183-84
- thinking. *See* moral reasoning
- three-step test (universalizability principle), 149-50, 164, 166
- Tiananmen Square, 188
- toddlers. *See* infants
- tolerance argument, 310-12
- torture
- conscientiousness and, 59
 - ethical pluralism and, 203-6
 - God and, 63, 64, 69
 - immorality and, 203-5
 - Mohammed and, 203-4
 - sadist and, 20, 114
 - Soviet psychiatric abuse and, 32-33, 37n4
 - of terrorists, 204-6
 - torture is immoral (example), 298-299, 300, 301, 302
 - waterboarding, 204
- totalitarian regimes, 135
- tragic dilemmas* argument, 250-51
- traits. *See* human traits
- trajectory* argument, 34-35
- Treatise on Human Nature* (Hume), 85n1
- trolley example, 217
- truth(s)
- conceptual, 73-74
 - empirical, 73, 74
 - logic and, 9
 - objective truths* argument, 307
 - subjectivity of, 307
- two worlds* argument, 29-30, 37
- unconditional value, 20, 163. *See also* good will
- happiness and, 20, 163
 - Kant and, 163
- understanding. *See* moral understanding
- unhappiness, 19. *See also* happiness; misery
- harm and, 35-36
- UNICEF, 132, 216
- unified moral theory
- ethics and, 13, 14, 235
 - feminist ethics and, 267-68, 270, 271
- universal human traits, 77-78
- universalizability principle, 147-50, 232, 271
- assessment of, 154-55
 - maxims and, 147-50
 - proceduralism and, 177
 - three-step test and, 149-50, 164, 166
- untrustworthiness, error theory and, 296-97
- utilitarianism, 112-43
- actions/intentions and, 118-19, 131-32
 - actual v. expected* results in, 117-18
 - animal experimentation and, 123, 124
 - attractions of, 119-24
 - benefit/harm ratio and, 125, 128
 - benevolence and, 112, 120, 130, 155, 265, 266
 - Bentham and, 120, 127
 - conflict resolution and, 121-22
 - consequentialism and, 115
 - DDE and, 208-9
 - decision procedure and, 232

- utilitarianism (*continued*)
 decision procedure/rightness standard and, 130-31
 defined, 115
 demanding nature of, 129-32
 difficulties in, 125-43
 ethical monism and, 202
 ethical pluralism and, 207-8
 hedonism and, 115-16
 impartiality and, 119-20, 133-34, 143n2
 intrinsic value and, 114, 115, 123, 124, 134-35
 Kantian perspective *v.*, 163, 174-75
 killing innocent human beings and, 118, 120, 122, 126, 135-36, 138, 139, 142
maximizing goodness principle and, 115-16
 Mill and, 115, 121-22, 127, 129
 moral common sense and, 120-21
 moral community and, 123-24, 143, 175
 moral flexibility and, 122-23
 moral knowledge and, 116-17
 motivation and, 129-31
 prima facie duties and, 224-25
Principles of Morals and Legislation, 120
 welfare measurement and, 125-28
Utilitarianism (Mill), 121-22
 utility principle, 115, 122, 131, 139, 271, 304n4
- V-2 rocket, 78
- validity
 logically valid arguments and, 7, 8, 28, 300
 test for, 8
- values. *See* intrinsic values; objective values
- value theory, 1, 2. *See also* desire satisfaction theory; good life; hedonism
- vegetarianism
meat eating argument and, 9-11
 Singer and, 120
- veil of ignorance, 184
- ventilator, pulling plug and, 218
- VE thesis, 241, 255. *See also* virtue ethics
- vicarious punishment, 136-37
- vices. *See* viciousness
- viciousness, 29-30, 37n2
- Vick, Michael, 83
- virtue, 246-47
 good life and, 248-49
 nature of, 246-47
 Plato and, 248
two worlds argument and, 29-30
- virtue ethics, 240-58
 Aristotle and, 241
 attractions of, 241-49, 258
 children and, 242, 243, 245-46, 249
 demanding nature of, 253
 divine command theory *v.*, 257
 moral complexity and, 242-43
 moral conflict and, 254-56
 moral education and, 245-46
 moral guidance and, 251-53
 moral understanding and, 243-44
 objections to, 249-58
 priority problem and, 256-58
 rape and, 256, 257
 right action standard and, 241-42
 role models for, 254, 256
tragic dilemmas argument and, 250-51
 VE thesis in, 241, 255
- waterboarding, 204. *See also* torture
- Waugh, Alexander, 143n2
- welfare. *See* well-being
- well-being (welfare), 18. *See also* desire satisfaction theory; happiness; hedonism
 measurement, utilitarianism and, 125-28
 objective theories of, 40-41, 43, 44, 54, 55
 personal authority and, 22
- Wesley, John, 112, 131. *See also* consequentialism
- what if everyone did that?* test, 145, 146, 149, 155, 232. *See also* golden rule
- "What Is Wrong with Slavery" (Hare), 158-59, 175n1
- why be moral?*
 contractarianism/Hobbes and, 190-94
 ethical egoism and, 101-4, 106-8
- William of Occam, 319. *See also* Occam's razor
- The Wire*, 201n1
- Wisconsin v. Yoder*, 286-87
- women. *See also* feminist ethics; rape
 Aquinas on, 259, 273n2
 Aristotle on, 259, 273n1
 experiences of, feminist ethics and, 260, 262-65
 Gilligan and, 261-62, 262, 272, 273n5
 Kant on, 259, 273n3
 Kohlberg's six-stage model and, 261
 Leviticus and, 68
 moral reasoning and, 260, 261-62, 277
 rape vulnerability and, 262, 263
 Rousseau on, 259, 273n4
 as second-class citizens, 50
 sports and, 269
 ● *on the Subjection of Women* and, 120
- World Trade Center attacks, 203-4
- Yoder, Wisconsin v.*, 286-87
- You Cannot Be Serious* (McEnroe), 49, 55n1



"*The Fundamentals of Ethics* is a superb book: accessible, illuminating, and comprehensive. Through engaging questions and lively thought experiments, keeping jargon and overt instruction to an absolute minimum, Shafer-Landau leads the reader to an articulate and self-aware grasp of the fundamental issues. In sum: the best such introductory text I know of."

—Niko Kolodny, *University of California, Berkeley*

"This is a great introduction to ethics, the finest that I am aware of. I am eager to use it in my classes. . . . It is similar in structure to the wildly popular Rachels introductory ethics text but better. . . . Shafer-Landau's writing is perfect for his intended audience."

—David Sobel, *University of Nebraska, Lincoln*

"The main strength in my mind is the clarity of the writing. Shafer-Landau has a gift for conveying the core ideas of different ethical theories, as well as fairly presenting their various strengths and weaknesses. Students will come away from reading *The Fundamentals of Ethics* with a very good understanding of philosophical thinking about a wide variety of issues in ethical theory."

—Christian Miller, *Wake Forest University*

In *The Fundamentals of Ethics*, author Russ Shafer-Landau employs a uniquely engaging writing style to introduce students to the essential ideas of moral philosophy. Offering more comprehensive coverage of the good life, normative ethics, and metaethics than any other text of its kind, this book also addresses issues that are often omitted from other texts, such as the doctrine of doing and allowing, the doctrine of double effect, ethical particularism, the desire-satisfaction theory of well-being, and moral error theory. Shafer-Landau carefully reconstructs and analyzes dozens of arguments in depth, at a level that is understandable to students with no prior philosophical background. Ideal for courses in introductory ethics and contemporary moral problems, this book can be used as a stand-alone text or with the author's companion reader, *The Ethical Life: Fundamental Readings in Ethics and Moral Problems*, which offers original readings exploring the topics covered in *The Fundamentals of Ethics*.

ABOUT THE AUTHOR

RUSS SHAFER-LANDAU is Professor of Philosophy at the University of Wisconsin, Madison, and the author of *Whatever Happened to Good and Evil?* (OUP, 2003) and *Moral Realism: A Defence* (OUP, 2003). He is also the editor of *The Ethical Life* (OUP, 2009), which is designed to be used along with this textbook.

Cover design by Black Kat Design

Cover painting: Singed Sumac and New Grass by Paul Horvath.
www.paulhorvath.com

OXFORD
UNIVERSITY PRESS

www.oup.com/us/he

ISBN 978-0-19-532086-2



9 780195 320862