

Are There Absolute Moral Rules?

You may not do evil that good may come.

SAINT PAUL, *LETTER TO THE ROMANS* (ca. A.D. 50)

9.1. Harry Truman and Elizabeth Anscombe

Harry S. Truman will always be remembered as the man who made the decision to drop the atomic bombs on Hiroshima and Nagasaki. When he became president in 1945, following the death of Franklin D. Roosevelt, Truman knew nothing about the bomb; Roosevelt's advisors had to fill him in. The Allies were winning the war in the Pacific, they said, but at a terrible cost. Plans had been drawn up for an invasion of Japan, but that battle would be even bloodier than the D-Day assault on Normandy, France, had been. Using the atomic bomb on one or two Japanese cities might bring the war to a speedy end, making the invasion unnecessary.

At first Truman was reluctant to use the new weapon. The problem was that each bomb would obliterate an entire city—not just the military targets, but the hospitals, schools, and homes. Women, children, old people, and other non-combatants would be wiped out along with the military personnel. The Allies had bombed cities before, but Truman sensed that the new weapon made the issue of noncombatant deaths more acute. Moreover, the United States was on record as condemning attacks on civilian targets. In 1939, before America had entered the war, President Roosevelt had sent a message to the governments of France, Germany, Italy, Poland, and Great Britain, denouncing the bombardment of

cities in the strongest terms. He had called it an “inhuman barbarism”:

The ruthless bombing from the air of civilians . . . which has resulted in the maiming and in the death of thousands of defenseless men, women, and children, has sickened the hearts of every civilized man and woman, and has profoundly shocked the conscience of humanity. If resort is had to this form of inhuman barbarism during the period of the tragic conflagration with which the world is now confronted, hundreds of thousands of innocent human beings who have no responsibility for, and who are not even remotely participating in, the hostilities which have now broken out, will lose their lives.

Truman expressed similar thoughts when he decided to authorize the bombings. He wrote in his diary, “I have told the Sec. of War, Mr. Stimson, to use it so that military objectives and soldiers and sailors are the target and not women and children. . . . The target will be a purely military one.” It is hard to know what to make of this, because Truman knew that the bombs would destroy whole cities. Nonetheless, it is clear that he was worried about the issue of noncombatants.

It is also clear that Truman was sure of his decision. Winston Churchill, the wartime leader of Great Britain, met with Truman shortly before the bombs were dropped. “The decision whether or not to use the atomic bomb,” Churchill later wrote, “. . . was never even an issue. There was unanimous, automatic, unquestioned agreement around our table.” Truman said that he “slept like a baby” after signing the final order.

Elizabeth Anscombe, who died in 2001, was a 20-year-old student at Oxford University when World War II began. At that time she co-authored a pamphlet arguing that Britain should not go to war because countries at war inevitably end up fighting by unjust means. “Miss Anscombe,” as she was always known—despite her 59-year marriage and her seven children—would go on to become one of the 20th century’s most distinguished philosophers and perhaps the greatest woman philosopher in history.

Miss Anscombe was also a Catholic, and her religion was central to her life. Her ethical views reflected traditional

Catholic teachings. In 1968, after Pope Paul VI affirmed the church's ban on contraception, she wrote a pamphlet explaining why artificial birth control is immoral. Late in her life, she was arrested while protesting outside a British abortion clinic. She also accepted the church's teaching about the ethical conduct of war, which brought her into conflict with Truman.

Harry Truman and Elizabeth Anscombe crossed paths in 1956. Oxford University was planning to grant Truman an honorary degree in thanks for America's wartime help, and those proposing the honor assumed that it would be uncontroversial. But Anscombe and two other faculty members opposed the idea. Although they lost, they forced a vote on what would otherwise have been a rubber-stamp approval. Then, while the degree was being conferred, Anscombe knelt outside the hall, praying.

Anscombe wrote another pamphlet, this time explaining that Truman was a murderer because he had ordered the bombings of Hiroshima and Nagasaki. Truman, of course, thought the bombings were justified because they had shortened the war and saved lives. For Anscombe, this was not good enough. "For men to choose to kill the innocent as a means to their ends," she wrote, "is always murder." To the argument that the bombings saved more lives than they took, she replied, "Come now: if you had to choose between boiling one baby and letting some frightful disaster befall a thousand people—or a million people, if a thousand is not enough—what would you do?"

Anscombe's example was apt. The bomb blast at Hiroshima, which ignited birds in midair, did lead to babies being boiled: People died in rivers, reservoirs, and cisterns, trying in vain to escape the heat. Anscombe's point was that *some things may not be done, no matter what*. It does not matter if we could accomplish some great good by boiling a baby; it is simply wrong. Anscombe believed in a host of such rules. Under no circumstances, she said, may we intentionally kill innocent people; worship idols; make a false profession of faith; engage in sodomy or adultery; punish one person for the acts of another; or commit treachery, which she describes as "obtaining a man's confidence in a grave matter by promises of trustworthy friendship and then betraying him to his enemies."

Anscombe's husband, Peter Geach (1916–2013), agreed with this. Anscombe and Geach were the 20th century's foremost philosophical champions of the doctrine that moral rules are absolute.

9.2. The Categorical Imperative

The idea that moral rules have no exceptions is hard to defend. It is easy enough to explain why we sometimes *should* break a rule—we can simply point to cases in which following the rule would have terrible consequences. But how can we defend *not* breaking the rule in such cases? We might say that moral rules are God's inviolable commands. Apart from that, what can be said?

Before the 20th century, there was one major philosopher who believed that moral rules are absolute. Immanuel Kant (1724–1804) argued that lying is wrong under any circumstances. He did not appeal to religion; instead, he held that lying is forbidden by reason itself. To see how he reached this conclusion, let's look at his general theory of ethics.

Kant observed that the word *ought* is often used nonmorally:

- If you want to become a better chess player, you *ought* to study the games of Magnus Carlsen.
- If you want to go to college, you *ought* to take the SAT.

Much of our conduct is governed by such “oughts.” The pattern is this: We have a certain desire (to become a better chess player, to go to college); we recognize that a certain course of action will help us get what we want (studying Carlsen's games, taking the SAT); and so we follow the indicated plan.

Kant called these “hypothetical imperatives” because they tell us what to do *provided that* we have the relevant desires. A person who did not want to improve her chess would have no reason to study Carlsen's games; someone who did not want to go to college would have no reason to take the SAT. Because the binding force of the “ought” depends on having the relevant desire, we can escape its grip by letting go of the desire. So, for example, I can avoid taking the SAT by deciding that I don't want to go to college.

Moral obligations, by contrast, do not depend on having particular desires. The form of a moral obligation is not “If you want so-and-so, then you ought to do such-and-such.” Instead, moral requirements are *categorical*: They have the form “You ought to do such-and-such, *period*.” The moral rule is not, for example, that you ought to help people *if* you care about them or *if* you want to be a good person. Instead, the rule is that you should help people *no matter what* your desires are. That is why moral requirements cannot be escaped by saying “I don’t care about that.”

Hypothetical “oughts” are easy to understand. They merely tell us to do what is necessary to achieve our goals. Categorical “oughts,” on the other hand, are mysterious. How can we be obligated to behave in a certain way regardless of our goals? Kant has an answer. Just as hypothetical “oughts” are possible because we have *desires*, categorical “oughts” are possible because we have *reason*. Categorical oughts, Kant says, are derived from a principle that every rational person must accept: the Categorical Imperative. In his *Foundations of the Metaphysics of Morals* (1785), Kant expresses the Categorical Imperative as follows:

Act only according to that maxim by which you can at the same time will that it should become a universal law.

This principle provides a way to tell whether an act is morally allowed. When you are thinking about doing something, ask what rule you would be following if you actually did it. This rule will be the “maxim” of your act. Then ask whether you would be willing for your maxim to become a universal law. In other words, would you allow your rule to be followed by all people at all times? If so, then your maxim is sound, and your act is acceptable. But if not, then your act is forbidden.

Kant gives several examples of how this works. Suppose, he says, a man needs money, but no one will lend it to him unless he promises to pay it back—which he knows he won’t be able to do. Should he make a false promise to get the loan? If he did, his maxim would be: *Whenever you need a loan, promise to repay it, even if you know you can’t*. Now, could he will that this rule become a universal law? Obviously not, because

it would be self-defeating. Once this rule became a universal practice, no one would believe such promises, and so no one would make loans based on them.

Kant gives another example, about giving aid. Suppose, he says, I refuse to help others in need, saying to myself, "What do I care? Let each person fend for himself." This, again, is a rule that I cannot will to be a universal law. For at some point in my life, I will need the help of others, and I will not want them to turn away from me.

9.3. Kant's Arguments on Lying

According to Kant, then, our behavior should be guided by universal laws, which are moral rules that hold true in all circumstances. Kant believed in many such exceptionless rules. We'll focus on the rule against lying, which Kant had especially strong feelings about. He said that lying under any circumstances is "the obliteration of one's dignity as a human being."

Kant offered two arguments for an absolute or exceptionless rule against lying.

1. His main argument relies on the Categorical Imperative. We could not will a universal law that allows us to lie, Kant said, because such a law would be self-defeating. As soon as lying became common, people would stop believing each other. Lying would then have no point, and in a sense lying would become impossible, because nobody would pay attention to what is said. Therefore, Kant reasoned, lying cannot be allowed. And so, it is forbidden under any circumstances.

This argument has a flaw, which will become clearer with an example. Suppose it was necessary to lie in order to save someone's life. Should you do it? Kant would reason as follows:

- (1) We should do only those actions that conform to rules which we could will to be adopted universally.
- (2) If you were to lie, then you would be following the rule, "It is okay to lie."
- (3) This rule could not be adopted universally, because it would be self-defeating: People would stop believing one another, and then it would be impossible to lie.
- (4) Therefore, you should not lie.

Although Anscombe agreed with Kant's conclusion, she was quick to point out an error in his reasoning. The difficulty arises in step (2). Why should we say that, if you lied, you would be following the rule, "It is okay to lie"? Perhaps your maxim would be: "I will lie when doing so would save someone's life." *That* rule would not be self-defeating. It could become a universal law. And so, by Kant's own theory, it would be all right for you to lie. Thus, Kant's belief that lying is always wrong does not seem to be in line with his own moral theory.

2. Many of Kant's contemporaries thought that his insistence on absolute rules was strange. One reviewer challenged him with this example: Imagine that someone is fleeing from a murderer and tells you that he is going home to hide. Then the murderer comes by and asks you where the man is. You believe that, if you tell the truth, you will be aiding in a murder. Furthermore, the killer is already headed the right way, so if you simply remain silent, the worst result is likely. What should you do? Let's call this the Case of the Inquiring Murderer. Under these circumstances, most of us believe that you should lie. After all, which is more important: telling the truth or saving someone's life?

Kant responded in an essay with the charmingly old-fashioned title "On a Supposed Right to Lie from Altruistic Motives," in which he gives a second argument against lying. Perhaps, he says, the man on the run has actually left his home, and by telling the truth you would lead the killer to look in the wrong place. However, if you lie, the murderer may wander away and discover the man leaving the area, in which case you would be responsible for his death. Whoever lies, Kant says, "must answer for the consequences, however unforeseeable they were, and pay the penalty for them." Kant states his conclusion in the tone of a stern schoolmaster: "To be truthful . . . in all declarations, therefore, is a sacred and absolutely commanding decree of reason, limited by no expediency."

This argument may be stated in a general form: We are tempted to make exceptions to the rule against lying because in some cases we think that the consequences of honesty will be bad and that the consequences of lying will be good. However, we can never be certain about what the consequences will be—we cannot *know* that good results will follow. The results of lying might be unexpectedly bad. Therefore, the best policy

is to avoid the known evil—lying—and let the consequences come as they may. Even if the consequences are bad, they will not be our fault, for we will have done our duty.

A similar argument would apply to Truman's decision to drop atomic bombs on Hiroshima and Nagasaki. The bombs were dropped in the hope that it would end the war. But Truman did not know that this would happen. The Japanese might have hunkered down, and the invasion might still have been necessary. So, Truman was betting hundreds of thousands of lives on the mere hope that good results might ensue.

The problems with this argument are obvious enough—so obvious, in fact, that it is surprising that a philosopher of Kant's caliber was not more sensitive to them. In the first place, the argument depends on an unreasonably pessimistic view of what we can know. Sometimes we can be quite confident of what the consequences will be, in which case we need not hesitate because of uncertainty. Moreover—and this is more significant, philosophically—Kant seems to assume that we would be morally responsible for any bad consequences of lying, but we would *not* be responsible for any bad consequences of telling the truth. Suppose, as a result of our telling the truth, the murderer found his victim and killed him. Kant seems to assume that we would be blameless. But can we escape responsibility so easily? After all, we told the murderer where to go. This argument, then, is not convincing.

Thus, Kant has failed to prove that lying is always wrong. The Case of the Inquiring Murderer shows what a tough row he chose to hoe. While Kant believes that lying “obliterates one's dignity as a human being,” common sense says that some lies are harmless. In fact, we have a name for them: white lies. Aren't white lies acceptable—or even required—when they can be used to save someone's life? This points to the main difficulty for the belief in absolute rules: Shouldn't a rule be broken when following it would be disastrous?

9.4. Conflicts between Rules

Suppose it is held to be absolutely wrong to do X in any circumstances and also wrong to do Y in any circumstances. Then what about the case in which a person must choose

between doing X and doing Y? This kind of conflict seems to show that moral rules can't be absolute.

Is there any way that this objection can be met? One way is to deny that such conflicts ever occur. Peter Geach took this view, appealing to God's providence. We can describe fictitious cases in which there is no way to avoid violating one of the absolute rules, he said, but God will not permit such circumstances to arise. Geach writes:

If God is rational, he does not command the impossible; if God governs all events by his providence, he can see to it that circumstances in which a man is inculpably faced by a choice between forbidden acts do not occur. Of course such circumstances . . . are consistently describable; but God's providence could ensure that they do not in fact arise. Contrary to what nonbelievers often say, belief in the existence of God does make a difference to what one expects to happen.

Do such cases actually occur? There is no doubt that serious moral rules sometimes clash. During World War II, Dutch fishermen smuggled Jewish refugees to England in their boats, and sometimes they would be stopped by Nazi patrols. The Nazi captain would call out and ask the Dutch captain where he was going, who was on board, and so forth. The fishermen would lie and be allowed to pass. Clearly, the fishermen had only two options: either lie or let everyone on board be killed. No third alternative was available; they could not, for example, remain silent or outrun the Nazis. Thus, Geach appears to have been naïve. Terrible dilemmas do occur in the real world.

If such dilemmas occur, then doesn't this disprove the existence of absolute moral rules? Suppose, for example, the two rules "It is wrong to lie" and "It is wrong to facilitate the murder of innocent people" are both taken to be absolute. The Dutch fishermen would have to do one of these things; therefore, a moral view that absolutely prohibits both is incoherent.

This type of argument is impressive, but it is also limited. It can be levied only against *pairs* of absolute moral rules; two rules are needed to create the conflict. The argument won't stop someone from believing that there is just one absolute

rule. And, in a way, everyone does believe in such a rule. “Do what is right” is a moral principle we all believe in, which admits of no exceptions. We should always do what is right. However, this rule is so formal that it is trivial—we believe it because it doesn’t really say anything. It is not the kind of absolute moral rule that Kant, Geach, and Anscombe wanted to argue for.

9.5. Kant’s Insight

Few contemporary philosophers would defend Kant’s Categorical Imperative. Yet it might be wrong to dismiss that principle too quickly. As Alasdair MacIntyre (1929–) observes, “For many who have never heard of philosophy, let alone of Kant, morality is roughly what Kant said it was”—that is, a system of rules that one must follow from a sense of duty. Is there some basic idea underlying the Categorical Imperative that we might accept, even if we don’t believe in absolute moral rules? I think there is.

Remember that Kant viewed the Categorical Imperative as binding on rational agents simply because they are rational; in other words, a person who rejected this principle would be guilty not merely of being immoral but also of being irrational. This is a compelling idea. But what exactly does it mean? In what sense would it be irrational to reject the Categorical Imperative?

Note that a moral judgment must be backed by good reasons—if it is true that you ought (or ought not) to do such-and-such, then there must be a reason why you should (or should not) do it. For example, you might think that you ought not to set forest fires because property would be destroyed and people would be killed. The Kantian twist is to point out that *if you accept any considerations as reasons in one case, then you must accept them as reasons in other cases as well*. If there is another case in which property would be destroyed and people killed, you must accept this as a reason in that case, too. It is no good saying that you can accept reasons some of the time, but not all the time; or that other people must respect them, but not you. Moral reasons, if they are valid at all, are binding on all people at all times. This is a requirement of consistency, and Kant was right to think that no rational person may deny it.

This insight has some important implications. It implies that a person cannot regard herself as special, from a moral point of view: She cannot consistently think that she is permitted to act in ways that are forbidden to others, or that her interests are more important than other people's interests. As one person put it, I cannot say that it is all right for me to drink your beer and then complain when you drink mine. If Kant was not the first to recognize this, he was the first to make it the cornerstone of a fully worked-out system of morals.

But Kant went one step further and said that consistency requires rules that have no exceptions. One can see how his insight pushed him in that direction; but the extra step was not necessary, and it causes trouble for his theory. Rules, even within a Kantian framework, need not be absolute. All that Kant's basic idea requires is that when we violate a rule, we do so for a reason that we would be willing for anyone to accept. In the Case of the Inquiring Murderer, this means that we may violate the rule against lying only if we would be willing for anyone else to lie in the same circumstances. And most of us would readily agree to that.

President Truman could also say that anyone in his position would have been justified in dropping the bomb. Thus, even if Truman was wrong, Kant's arguments do not prove it. One might say that dropping the bomb was wrong because Truman had better options. Perhaps he should have shown the Japanese the power of the bomb by dropping it onto an unpopulated area—negotiations might then have been successful. Or perhaps the Allies could have simply declared victory at that point in the war, even without a Japanese surrender. Saying *things like that*, however, is very different from saying that what Truman did violated an absolute rule.

CHAPTER 10

Kant and Respect for Persons

Are there any who would not admire man?

GIOVANNI PICO DELLA MIRANDOLA,
ORATION ON THE DIGNITY OF MAN (1486)

10.1. Kant's Core Ideas

Immanuel Kant thought that human beings occupy a special place in creation. Of course, he was not alone in thinking this. From ancient times, humans have considered themselves to be essentially different from all other creatures—and not just different, but better. In fact, humans have traditionally considered themselves to be quite fabulous. Kant certainly did. On his view, human beings have “an intrinsic worth” or “dignity” that makes them valuable “above all price.”

Other animals, Kant thought, have value only insofar as they serve human purposes. In his *Lectures on Ethics* (1779), Kant writes, “But so far as animals are concerned, we have no direct duties. Animals . . . are there merely as means to an end. That end is man.” We may, therefore, use animals in any way we please. We don’t even have a “direct duty” to refrain from torturing them. Kant did condemn the abuse of animals, but not because the animals would be hurt. He worried, rather, about us: “He who is cruel to animals also becomes hard in his dealings with men.”

When Kant said that human beings are valuable “above all price,” this was not mere rhetoric. Kant meant that people are irreplaceable. If a child dies, this is a tragedy, and it remains tragic even if another child is born into the same family. On the other hand, “mere things” are replaceable. If your printer

breaks, then everything is fine if you can get another printer. People, Kant believed, have a “dignity” that mere things lack.

Two facts about people, Kant thought, support this judgment.

First, because people have desires, things that satisfy those desires can have value *for* people. By contrast, “mere things” have value only insofar as they promote human ends. Thus, if you want to become a better poker player, a book about poker will have value for you; but, apart from such ends, books about poker are worthless. Or, if you want to go somewhere, a car will have value for you; yet apart from such desires, cars have no value.

Mere animals, Kant thought, are too primitive to have desires and goals. Thus, they are “mere things.” Kant did not believe, for example, that milk has value *for* the cat who wishes to drink it. Today, however, we are more impressed with the mental life of animals than Kant was. We believe that nonhuman animals do have desires and goals. So, perhaps there are Kantian grounds for saying that animals are not “mere things.”

Yet Kant’s second reason would not apply to animals. People, Kant said, have “an intrinsic worth, i.e., dignity” because they are *rational agents*, that is, free agents capable of making their own decisions, setting their own goals, and guiding their conduct by reason. The only way that moral goodness can exist, Kant held, is for rational creatures *to act from a good will*—that is, to apprehend what they should do and act from a sense of duty. Human beings are the only rational agents that exist on earth; nonhuman animals lack free will, and they do not “guide their conduct by reason” because their rational capacities are too limited. If people disappeared, then so would the moral dimension of the world. This second fact is especially important for Kant.

Thus, Kant believed, human beings are not merely one valuable thing among others. Humans are the ones who do the valuing, and it is their conscientious actions alone that have moral worth. Human beings tower above the realm of things.

These thoughts are central to Kant’s moral system. Kant believed that all of our duties can be derived from one ultimate principle, which he called the Categorical Imperative.

Kant gave this principle different formulations, but at one point he put it like this:

Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only.

Because people are so valuable, morality requires us to treat them “always as an end and never as a means only.” What does this mean, and why should anyone believe it?

To treat people “as an end” means, on the most superficial level, treating them well. We must promote their welfare, respect their rights, avoid harming them, and generally “endeavor, so far as we can, to further the ends of others.” But Kant’s idea also has a deeper implication. To treat people as ends requires treating them with respect. Thus, we may not manipulate people or “use” people to achieve our goals, no matter how good those goals may be. Kant gives this example: Suppose you need money, and you want a loan, but you know you cannot repay it. In desperation, you consider telling your friend you will repay it in order to get the money. May you do this? Perhaps you need the money for a good purpose—so good, in fact, that you might convince yourself that a lie would be justified. Nevertheless, you should not lie to your friend. If you did, you would be manipulating her and using her “merely as a means.”

On the other hand, what would it be like to treat your friend “as an end”? Suppose you tell the truth—you tell her why you need the money, and you tell her you won’t be able to pay her back. Then your friend can make up her own mind about whether to give you the loan. She can consult her own values and wishes, exercise her own powers of reasoning, and make a free choice. If she then decides to give you the money for your stated purpose, she will be choosing *to make that purpose her own*. Thus, you will not be using her as a mere means to achieving your goal, for it will be her goal, too. Thus, for Kant, to treat people as ends is to treat them “as beings who [can] contain in themselves the end of the very same action.”

When you tell your friend the truth, and she gives you money, you are using her as a means to getting the money. However, Kant does not object to treating someone as a means; he objects to treating someone *only* as a means. Consider

another example: Suppose your bathroom sink is stopped up. Would it be okay to call in a plumber—to “use” the plumber as a means to unclogging the drain? Kant would have no problem with this, either. The plumber, after all, understands the situation. You are not deceiving or manipulating him. He may freely choose to unclog your drain in exchange for payment. Although you are treating the plumber as a means, you are also treating him with dignity, as an “end-in-himself.”

Treating people as ends, and respecting their rational capacities, has other implications. We should not force adults to do things against their will; instead, we should let them make their own decisions. We should therefore be wary of laws that aim to protect people from themselves—for example, laws requiring people to wear seat belts or motorcycle helmets. Also, we shouldn’t forget that respecting *people* requires respecting *ourselves*. I should take good care of myself; I should develop my talents; I should do more than just slide by.

Kant’s moral system is not easy to grasp. To understand it better, let’s consider how Kant applied his ideas to the practice of criminal punishment. The rest of this chapter is devoted to that example.

10.2. Retribution and Utility in the Theory of Punishment

Jeremy Bentham (1748–1832) said that “all punishment is mischief: all punishment in itself is evil.” Bentham had a point. As a society, we punish people by making them pay fines or go to prison. Sometimes we even kill them. Punishment, by its nature, always involves inflicting harm. How can it be right to hurt people?

The traditional answer is that punishment is justified as a way of “paying back” the offender for his wicked deed. Those who have committed a crime deserve to be treated badly. It is a matter of justice: If you harm other people, justice requires that you be harmed, too. As the ancient saying has it, “An eye for an eye, and a tooth for a tooth.” According to the doctrine of Retributivism, this is the main justification of punishment.

On Bentham’s view, Retributivism is a wholly unsatisfactory idea, because it advocates the infliction of suffering without any compensating gain in happiness. Retributivism

would have us increase, not decrease, the amount of misery in the world. Kant was a retributivist, and he openly embraced this implication. In *The Critique of Practical Reason* (1788), he writes:

When, however, someone who delights in annoying and vexing peace-loving folk receives at last a right good beating, it is certainly an ill, but everyone approves of it and considers it as good in itself even if nothing further results from it.

Thus, punishing people may increase the amount of misery in the world; but that is all right, for the extra suffering is borne by those who deserve it.

Utilitarianism takes a very different approach. According to Utilitarianism, our duty is to do whatever will increase the amount of happiness in the world. Punishment is, on its face, "an evil" because it makes the punished person unhappy. Thus, Bentham, a utilitarian, says, "If [punishment] ought at all to be admitted, it ought to be admitted in as far as it promises to exclude some greater evil." In other words, punishment can be justified only if it does enough good to outweigh the bad. And utilitarians have traditionally thought that it does. If someone breaks the law, then punishing that person can have several benefits.

First, punishment provides comfort and gratification to victims and their families. People feel very strongly that someone who mugged, raped, or robbed them should not go free. Victims also live in fear when they know that their attacker is still on the street. Philosophers usually ignore this justification of punishment, but it plays a prominent role in our legal system. Judges, lawyers, and juries often want to know what victims want. Indeed, whether the police will make an arrest, and whether the district attorney's office will prosecute a case, often depends on the wishes of the victims.

Second, by locking up criminals, or by executing them, we take them off the street. With fewer criminals on the loose, there is less crime. In this way, prisons protect society and thus reduce unhappiness. Of course, this justification does not apply to punishments in which the offender remains free, such as when a criminal is sentenced to probation with community service.

Third, punishment reduces crime by deterring would-be criminals. Someone who is tempted to commit a crime might not do so if he knows he might be punished. Obviously, the threat of punishment is not always effective; sometimes people break the law anyway. But there will be *less* misconduct if punishments are threatened. Imagine what would happen if the police stopped arresting thieves; surely there would be a lot more theft. Deterring crime thus prevents unhappiness.

Fourth, a well-designed system of punishment might help to rehabilitate wrongdoers. Criminals often have mental and emotional problems. Many are uneducated and illiterate and cannot hold down jobs. Why not respond to crime by attacking the problems that cause it? If someone is dangerous, we may imprison him. But while we have him behind bars, why not address his problems with psychological therapy, educational opportunities, and job training? If one day he can return to society as a productive citizen, then both he and society will benefit.

In America, the utilitarian view of punishment was once dominant. In 1954, the American Prison Association changed its name to "the American Correctional Association" and encouraged prisons to become "correctional facilities." Prisons were thus asked to "correct" inmates, not to "punish" them. Prison reform was common in the 1950s and 1960s. Prisons offered their inmates drug treatment programs, vocational training classes, and group counseling sessions, hoping to turn them into good citizens.

Those days, however, are long gone. In the 1970s, the newly declared "war on drugs" led to longer and longer prison sentences for drug offenders. This change in American justice was more retributive than utilitarian in nature, and it resulted in vastly more prisoners at any given time. Today the United States houses around 2.3 million inmates, giving it the highest incarceration rate of any country, by far. Most of those inmates are in state prisons, not federal prisons, and the states that must operate those facilities are strapped for cash. As a result, most of the programs aimed at rehabilitation were either scaled back or eliminated. The rehabilitation mentality of the 1960s has thus been replaced by a warehousing mentality, marked by prison overcrowding and plagued by underfunding. This new reality, which is less pleasant for the inmates themselves, suggests a victory for Retributivism.

10.3. Kant's Retributivism

The utilitarian theory of punishment has many opponents. Some critics say that prison reform does not work. California had the most vigorous program of reform in the United States, yet its prisoners were especially likely to commit crimes after being released. Most of the opposition, however, is based on theoretical considerations that go back at least to Kant.

Kant despised "the serpent-windings of Utilitarianism" because, he said, the theory is incompatible with human dignity. In the first place, it has us calculating how to use people as means to our ends. If we imprison the criminal in order to keep society safe, we are merely using him for the benefit of others. This violates Kant's belief that "one man ought never to be dealt with merely as a means subservient to the purpose of another."

Moreover, rehabilitation is really just the attempt to mold people into what *we* want them to be. As such, it violates their right to decide for themselves what sort of people they will be. We do have the right to respond to their wickedness by "paying them back" for it, but we do not have the right to violate their integrity by trying to manipulate their personalities.

Thus, Kant would have no part of utilitarian justifications. Instead, he believes that punishment should be governed by two principles. First, people should be punished simply because they have committed crimes, and for no other reason. Second, punishment should be *proportionate* to the seriousness of the crime. Small punishments may suffice for small crimes, but big punishments are necessary for big crimes:

But what is the mode and measure of punishment which public justice takes as its principle and standard? It is just the principle of equality, by which the pointer of the scale of justice is made to incline no more to the one side than to the other. . . . Hence it may be said: "If you slander another, you slander yourself; if you steal from another, you steal from yourself; if you strike another, you strike yourself; if you kill another, you kill yourself." This is . . . the only principle which . . . can definitely assign both the quality and the quantity of a just penalty.

Kant's second principle leads him to endorse capital punishment; for in response to murder, only death is appropriate. In a famous passage, Kant says:

Even if a civil society resolved to dissolve itself with the consent of all its members—as might be supposed in the case of a people inhabiting an island resolving to separate and scatter throughout the whole world—the last murderer lying in prison ought to be executed before the resolution was carried out. This ought to be done in order that everyone may realize the desert of his deeds, and that blood-guiltiness may not remain on the people; for otherwise they will all be regarded as participants in the murder. . . .

Although a Kantian must support the death penalty *in theory*, she might oppose it *in practice*. The worry, in practice, is that innocent people might be killed by mistake. In the United States, around 130 death row inmates have been released from prison after being proved innocent. None of those people were actually killed. But with so many close calls, it is almost certain that some innocent people have been put to death—and advocates of reform point to specific, troubling cases. Thus, in deciding whether to support a policy of capital punishment, Kantians must balance the injustice of the occasional, deadly mistake against the injustice of letting killers live.

Kant's principles describe a general theory of punishment: Wrongdoers must be punished, and the punishment must fit the crime. This theory is deeply opposed to the Christian idea of turning the other cheek. In the Sermon on the Mount, Jesus avows, "You have heard that it was said, 'An eye for an eye and a tooth for a tooth.' But I say to you, Do not resist the one who is evil. If anyone slaps you on the right cheek, turn to him the other also." For Kant, such a response to evil is not only imprudent, but unjust.

What arguments can be given for Kant's Retributivism? As we noted, Kant regards punishment as a matter of justice. He says that justice is not done if the guilty go unpunished. That is one argument. Also, we discussed why Kant rejects the utilitarian view of punishment. But he also gives another argument, based on his idea of treating people as "ends-in-themselves."

How does this argument go? On the face of it, it seems unlikely that we could describe punishing someone as "respecting him as a person" or as "treating him as an end." How could sending someone to prison be a way of respecting him? Even more paradoxically, how could executing someone be a way of treating him with dignity?

For Kant, treating someone "as an end" means treating him as a rational being, who is responsible for his behavior. So now we may ask: What does it mean to be a responsible being?

Let's first consider what it means *not* to be such a being. Mere animals, who lack reason, are not responsible for their actions; nor are people who are mentally ill and cannot control themselves. In such cases, it would be absurd to "hold them accountable." We could not properly feel gratitude or resentment toward them, because they are not responsible for any good or ill they cause. Moreover, we cannot expect them to understand *why* we treat them as we do, any more than they understand why they behave as they do. So we have no choice but to deal with them by manipulating them. When we scold a dog for eating off the table, for example, we are merely trying to "train" him.

On the other hand, a rational being can freely decide what to do, based on his own conception of what is best. Rational beings *are* responsible for their behavior, and so they are accountable for what they do. We may feel gratitude when they behave well and resentment when they behave badly. Reward and punishment—not "training" or other forms of manipulation—are the natural expressions of gratitude and resentment. Thus, in punishing people, we are holding them responsible for their actions in a way in which we cannot hold mere animals responsible. We are responding to them not as people who are "sick" or who have no control over themselves, but as people who have freely chosen their evil deeds.

Furthermore, in dealing with responsible agents, we may properly allow their conduct to determine, at least in part, how we respond to them. If someone has been kind to you, you may respond by being generous; and if someone is nasty to you, then you may take that into account in deciding how to respond. And why shouldn't you? Why should you treat everyone alike, regardless of how *they* have chosen to behave?

Kant gives this last point a distinctive twist. There is, on his view, a deep reason for responding to other people “in kind.” When we choose to do something, after consulting our own values, we are in effect saying *this is the sort of thing that should be done*. In Kant’s terminology, we are implying that our conduct be made into a “universal law.” Therefore, when a rational being decides to treat people in a certain way, he decrees that *this is the way people are to be treated*. Thus, if we treat him the same way in return, we are doing nothing more than treating him *as he has decided that people are to be treated*. If he treats others badly, and we treat him badly, we are complying with his own decision. We are, in a perfectly clear sense, respecting him by allowing his own judgment to control how we treat him. Of the criminal, Kant says, “His own evil deed draws the punishment upon himself.”

This last argument can be questioned. Why should we adopt the criminal’s principle of action, rather than our own? Shouldn’t we try to “be better than he is”? Also, bear in mind that even the wicked sometimes behave well. So if we treat the evildoer well, wouldn’t we also be following his judgment—a judgment that he has endorsed on many occasions?

At the end of the day, what we think of Kant’s theory may depend on our view of criminal behavior. If we see criminals as victims of circumstance, who do not ultimately control their own lives, then the utilitarian model will appeal to us. On the other hand, if we see criminals as rational agents who freely choose to do harm, then Kantian Retributivism will have more appeal. The resolution of this great debate about punishment might thus turn on whether we believe that human beings have free will, or whether we believe that outside forces impact human behavior so deeply that our freedom is an illusion. The debate about free will, however, is so complex, and so concerned with matters outside of ethics, that we will not discuss it here. This kind of dialectical situation is common in philosophy: when you study one matter deeply, you often come to realize that it depends on something else—something that is just as hard as the problem you began with.