



Kant and Deontological Theories

Kant's *Groundwork of the Metaphysics of Morals* is one of the small books which are truly great: it has exercised on human thought an influence almost ludicrously disproportionate to its size. In moral philosophy it ranks with the *Republic* of Plato and the *Ethics* of Aristotle; and perhaps it shows in some respects a deeper insight even than these. Its main topic—the supreme principle of morality—is of the utmost importance to all who are not indifferent to the struggle of good against evil. Its message was never more needed than it is at present, when a somewhat arid empiricism is the prevailing fashion in philosophy.

H. J. PATON, PREFACE TO KANT'S *GROUNDWORK OF THE METAPHYSIC OF MORALS*

Let's look again at our opening story in Chapter 7 on utilitarianism. A millionaire makes a dying request for you to donate \$2 million to the Yankees. You agree but then are tempted to give the money to the World Hunger Relief Organization instead. What should you do? The utilitarian, who focuses on the consequences of actions, would tell you to act in a way that advances the greatest good for the greatest number. In essence, the end justifies the means. Accordingly, breaking your promise to the millionaire and donating to the World Hunger Relief Organization appears to be the way to go.

The deontological answer to this question, however, is quite the opposite. It is not the consequences that determine the rightness or wrongness of an act but certain features in the act itself or in the rule of which the act is a token or example. The end never justifies the means. For example, there is something right about truth telling and promise keeping even when such actions may bring

about some harm; and there is something wrong about lying and promise breaking even when such actions may bring about good consequences. Acting unjustly is wrong even if it will maximize expected utility.

In this chapter, we explore deontological approaches of ethics, specifically that by Immanuel Kant (1724–1804). The greatest philosopher of the German Enlightenment and one of the most important philosophers of all time, Kant was both an absolutist and a rationalist. He believed that we could use reason to work out a consistent, nonoverridable set of moral principles.

KANT'S INFLUENCES

To understand Kant's moral philosophy, it is helpful to know a little about his life. He was born in Königsberg, Germany, in 1724, and died there eighty years later, never having left the surroundings of the city. He was a short, quiet man and so methodical that, as tradition has it, the citizens of Königsberg set their watches by his daily 3:00 p.m. walks. He never married. He devoted his life to the study and teaching of philosophy at the University of Königsberg. His greatest publication, *The Critique of Pure Reason* (1781), was heralded in his own day as a monumental work, and his *Foundations of the Metaphysics of Morals* (1785) is generally regarded as one of the two or three most important books in the history of ethics.

There were several strong influences on Kant's ethical thinking. The first was Pietism, a sect within the Lutheran Church, which Kant's parents adhered to and which set a tone of deep sincerity to his views. Pietists emphasized honesty, deep feeling, and the moral life rather than theological doctrine or orthodox belief. It's a religion of the heart, not the head, of the spirit rather than of ritual. However, as an intellectual, Kant emphasized the head as much as the heart, but it was a head concerned about the moral life, especially good will. It is not correct beliefs or results that really matter but inner goodness. The idea is that, if we live within our lights, we will be given more light and that God judges us not on how lucky or successful we are in accomplishing our tasks but on how earnestly we have lived according to our principles. This impacted Kant's notion of the good will as the sole intrinsic good in life.

The second influence was the work of Jean-Jacques Rousseau (1712–1778) on human freedom, especially his *Social Contract*, and it was said that the only time Kant ever missed his afternoon walk was the day when he read that book. Rousseau taught him the meaning and importance of human dignity, the primacy of freedom and autonomy, and the intrinsic worth of human beings apart from any functions they might perform.

The third influence was the philosophical debate of his time between **rationalism** and **empiricism**. The fourth influence was natural law intuitionist theories that dominated moral philosophy at that time. Let's look at these latter two influences in more detail.

Rationalism and Empiricism

The philosophical debate between rationalism and empiricism took place in the seventeenth and eighteenth centuries. Rationalists, such as René Descartes, Baruch Spinoza, Gottfried Leibniz, and Christian Wolff, claimed that pure reason could tell us how the world is, independent of experience. We can know metaphysical truth such as the existence of God, the immortality of the soul, freedom of the will, and the universality of causal relations apart from experience. Experience may be necessary to open our minds to these ideas, but essentially they are innate ideas that God implants in us from birth. Empiricists, led by John Locke and David Hume, on the other hand, denied that we have any innate ideas and argued that all knowledge comes from experience. Our minds are a *tabula rasa*, an empty slate, upon which experience writes her lessons.

The rationalists and empiricists carried their debate into the area of moral knowledge. The rationalists claimed that our knowledge of moral principles is a type of metaphysical knowledge, implanted in us by God, and discoverable by reason as it deduces general principles about human nature. On the other hand, empiricists, especially Francis Hutcheson, David Hume, and Adam Smith, argued that morality is founded entirely on the contingencies of human nature and based on desire. Morality concerns making people happy, fulfilling their reflected desires, and reason is just a practical means of helping them fulfill their desires. There is nothing of special importance in reason in its own right. It is mainly a rationalizer and servant of the passions. As Hume said, "Reason is and ought only to be a slave of the passions and can never pretend to any other office than to serve and obey them." Morality is founded on our feeling of sympathy with other people's sufferings, on fellow feeling. For such empiricists then, morality is contingent upon human nature:

Human nature → Feelings and Desires → Moral principles

If we had a different nature, then we would have different feelings and desires, and hence we would have different moral principles.

Kant rejected the ideas of Hutcheson, Hume, and Smith. He was outraged by the thought that morality should depend on human nature and be subject to the fortunes of change and the luck of empirical discovery. Morality is not contingent but necessary. It would be no less binding on us if our feelings were different from what they are. Kant writes,

Every empirical element is not only quite incapable of being an aid to the principle of morality, but is even highly prejudicial to the purity of morals; for the proper and inestimable worth of an absolutely good will consists just in this, that the principle of action is free from all influence of contingent grounds, which alone experience can furnish. We cannot too much or too often repeat our warning against this lax and even mean habit of thought which seeks for its principle amongst empirical motives and laws; for human reason in its weariness is glad to rest on this pillow, and in a dream of sweet illusions it substitutes for morality a

bastard patched up from limbs of various derivation, which looks like anything one chooses to see in it; only not like virtue to one who has once beheld her in her true form.¹

No, said Kant, it is not our desires that ground morality but our rational will. Reason is sufficient for establishing the moral law as something transcendent and universally binding on all rational creatures.

Act- and Rule-Intuitionism

Since the Middle Ages, one of the dominant versions of European moral philosophy was **natural law theory**. In a nutshell, this view maintained that, through rational intuitions embedded in human nature by God, we discover eternal and absolute moral principles. Medieval natural law philosopher Thomas Aquinas argued that we have a special mental process called *synderesis* that gives us general knowledge of moral goodness. From this knowledge, then, we derive a series of basic moral obligations. What's key here is the idea that humans have a natural faculty that gives us an intuitive awareness of morality. This general position is called **intuitionism**. During the seventeenth and eighteenth centuries, some sort of intuitionism was assumed in most ethical theories, and Kant was heavily influenced by some of them. Two basic forms emerged: act and rule intuitionism.

Act-intuitionism sees each act as a unique ethical occasion and holds that we must decide what is right or wrong in each situation by consulting our conscience or our intuitions or by making a choice apart from any rules. We must consult our conscience in every situation to *discover* the morally right (or wrong) thing to do. An expression of act-intuitionism is in the famous moral sermons of Joseph Butler (1692–1752), a bishop within the Church of England. He writes,

[If] any plain honest man, before he engages in any course of action, ask himself, Is this I am going about right, or is it wrong? . . . I do not in the least doubt but that this question would be answered agreeably to truth and virtue, by almost any fair man in almost any circumstance.²

Butler believed that we each have a conscience that can discover what is right and wrong in virtually every instance. This is consistent with advice such as “Let your conscience be your guide.” We do not need general rules to learn what is right and wrong; our intuition will inform us of those things. The judgment lies in the moral perception and not in some abstract, general rule.

Act-intuitionism, however, has some serious disadvantages. First, it is hard to see how any argument could take place with an intuitionist: Either you both have the same intuition about lying or you don't, and that's all there is to it. If I believe that a specific act of abortion is morally permissible and you believe it is morally wrong, then we may ask each other to look more deeply into our consciences, but we cannot argue about the subject. There *is* a place for deep intuitions in moral philosophy, but intuitions must still be scrutinized by reason and corrected by theory.

Second, it seems that rules are necessary to all reasoning, including moral reasoning, and act-intuitionists seem to ignore this. You may test this by thinking about how you learn to drive a car, to do long division, or to type. Even though you may eventually internalize the initial principles as habits so that you are unconscious of them, one could still cite a rule that covers your action. For example, you may no longer remember the rules for accelerating a car, but there was an original experience of learning the rule, which you continue unwittingly to follow. Moral rules such as “Keep your promises” and “Don’t kill innocent people” seem to function in a similar way.

Third, different situations seem to share common features, so it would be inconsistent for us to prescribe different moral actions. Suppose you believe that it is morally wrong for John to cheat on his math exam. If you also believe that it is morally permissible for you to cheat on the same exam, don’t you need to explain what makes your situation different from John’s? If I say that it is wrong for John to cheat on exams, am I not implying that it is wrong for anyone relevantly similar to John (including all students) to cheat on exams? That is, morality seems to involve a universal aspect, or what is called the principle of universalizability: If one judges that X is right (or wrong) or good (or bad), then one is rationally committed to judging anything relevantly similar to X as right (wrong) or good (bad). If this principle is sound, then act-intuitionism is misguided.

The other intuitionist approach, **rule-intuitionism**, maintains that we must decide what is right or wrong in each situation by consulting moral rules that we receive through intuition. Rule-intuitionists accept the principle of universalizability as well as the notion that in making moral judgments we are appealing to principles or rules. Such rules as “We ought never to lie,” “We ought always to keep our promises,” and “We ought never to execute an innocent person” constitute a set of valid prescriptions regardless of the outcomes. The rule-intuitionist to have the greatest impact on Kant was German philosopher Samuel Pufendorf (1632–1694), the dominant natural law theorist of his time. Pufendorf describes the intuitive process by which we acquire moral knowledge:

It is usually said that we have knowledge of this [moral] law from nature itself. However, this is not to be taken to mean that plain and distinct notions concerning what is to be done or avoided were implanted in the minds of newborn people. Instead, nature is said to teach us, partly because the knowledge of this law may be attained by the help of the light of reason. It is also partly because the general and most useful points of it are so plain and clear that, at first sight, they force assent. . . . Although we are not able to remember the precise time when they first took hold of our understandings and professed our minds, we can have no other opinion of our knowledge of this law except that it was native to our beings, or born together and at the same time with ourselves.³

The moral intuitions that we have, according to Pufendorf, fall into three groups: duties to God, to oneself, and to others. The duties in all these cases are moral

rules that guide our actions. Within these three groupings, the main rules of duty that Pufendorf advocates are these:

- *To God.* Know the existence and nature of God; worship God.
- *To oneself.* Develop one's skills and talents; avoid harming our bodies, such as through gluttony or drunkenness, and not killing oneself.
- *To others.* Avoid wronging others; treat people as equals; promote the good of others; keep one's promises.

Kant was influenced by Pufendorf in two ways. First, Kant was a rule-intuitionist of a special sort: He believed that moral knowledge comes to us through *rational* intuition in the form of moral rules. As we'll see, Kant's moral psychology is rather complex, and his conception of intuition draws on a distinct notion of reason, which we don't find in Pufendorf. Second, Kant accepted Pufendorf's division of duties toward God, oneself, and others. Duties toward God, Kant argues, are actually religious duties, not moral ones. However, duties to oneself and others are genuine moral obligations.

THE CATEGORICAL IMPERATIVE

The principal moral rule in Kant's ethical theory is what he calls the categorical imperative—essentially meaning “absolute command.” Before introducing us to the specific rule itself, he sets the stage with an account of intrinsic moral goodness.

Intrinsic Goodness and the Good Will

As we have noted, Kant wanted to remove moral truth from the zone of contingency and empirical observation and place it securely in the area of necessary, absolute, universal truth. Morality's value is not based on the fact that it has instrumental value, that it often secures nonmoral goods such as happiness; rather, morality is valuable in its own right:

Nothing can possibly be conceived in the world, or even out of it, which can be called good without qualification, except the Good Will. Intelligence, wit, judgment, and the other *talents* of the mind, however they may be named, or courage, resolution, perseverance, as qualities of temperament, as undoubtedly good and desirable in many respects; but these gifts of nature also may become extremely bad and mischievous if the will which is to make use of them, and which, therefore constitutes what is called *character* is not good. . . . Even if it should happen that, owing to special disfavor of fortune, or the stingy provision of a stepmotherly nature, this Good Will should wholly lack power to accomplish its purpose, if with its greatest efforts it should yet achieve nothing, and there should remain only the Good Will, . . . then, like a jewel, it

would still shine by its own light, as a thing which has its whole value in itself. Its usefulness or fruitfulness can neither add to nor take away anything from this value.⁴

The only thing that is absolutely good, good in itself and without qualification, is the good will. All other intrinsic goods, both intellectual and moral, can serve the vicious will and thus contribute to evil. They are only *morally valuable* if accompanied by a good will. Even success and happiness are not good in themselves. Honor can lead to pride. Happiness without good will is undeserved luck, ill-gotten gain. Nor is utilitarianism plausible, for if we have a quantity of happiness to distribute, is it just to distribute it equally, regardless of virtue? Should we not distribute it discriminately, according to moral goodness? Happiness should be distributed in proportion to people's moral worth.

How successful is Kant's argument for the good will? Could we imagine a world where people always and necessarily put nonmoral virtues to good use, where it is simply impossible to use a virtue such as intelligence for evil? Is happiness any less good simply because one can distribute it incorrectly? Can't one put the good will itself to bad use as the misguided do-gooder might? As the aphorism goes, "The road to hell is paved with good intentions." Could Hitler have had good intentions in carrying out his dastardly programs? Can't the good will have bad effects?

Although we may agree that the good will is a great good, it is not obvious that Kant's account is correct, that it is the only inherently good thing. For even as intelligence, courage, and happiness can be put to bad uses or have bad effects, so can the good will; and even as it doesn't seem to count against the good will that it can be put to bad uses, so it shouldn't count against the other virtues that they can be put to bad uses. The good will may be a necessary element to any morally good action, but whether the good will is also a *sufficient* condition to moral goodness is another question.

Nonetheless, perhaps we can reinterpret Kant so as to preserve his central insight. There does seem to be something morally valuable about the good will, apart from any consequences. Consider the following illustration. Two soldiers volunteer to cross enemy lines to contact their allies on the other side. Both start off and do their best to get through the enemy area. One succeeds; the other doesn't and is captured. But, aren't they both morally praiseworthy? The success of one in no way detracts from the goodness of the other. Judged from a commonsense moral point of view, their actions are equally good; judged from a utilitarian or consequentialist view, the successful act is far more valuable than the unsuccessful one. Here, we can distinguish the agent's worth from the value of the consequences and make two separate, nonconflicting judgments.

Hypothetical versus Categorical Imperatives

For Kant, all mention of duties (or obligations) can be translated into the language of imperatives, or commands. As such, moral duties can be said to have imperative force. He distinguishes two kinds of imperatives: hypothetical and categorical.

The formula for a **hypothetical imperative** is “If you want A, then do B.” For example, “If you want a good job, then get a good education,” or “If you want to be happy, then stay sober and live a balanced life.” The formula for a categorical imperative is simply: “Do B!” That is, do what reason discloses to be the intrinsically right thing to do, such as “Tell the truth!” Hypothetical, or means–ends, imperatives are not the kind of imperatives that characterize moral actions. Categorical, or unqualified, imperatives are the right kind of imperatives, because they show proper recognition of the imperial status of moral obligations. Such imperatives are intuitive, immediate, absolute injunctions that all rational agents understand by virtue of their rationality.

Kant argues that one must perform moral duty solely for its own sake (“duty for duty’s sake”). Some people conform to the moral law because they deem it in their own enlightened self-interest to be moral. But they are not truly moral because they do not act for the sake of the moral law. For example, a businessman may believe that “honesty is the best policy”; that is, he may judge that it is conducive to good business to give his customers correct change and high-quality products. But, unless he performs these acts *because* they are his duty, he is not acting morally, even though his acts are the same ones they would be if he *were* acting morally.

The kind of imperative that fits Kant’s scheme as a product of reason is one that universalizes principles of conduct. He names it the **categorical imperative** (CI): “Act only according to that maxim by which you can at the same time will that it would become a universal law.” The categorical imperative, for Kant, is a procedure for determining the morality of *any* course of action. All specific moral duties, he writes, “can be derived from this single imperative.” Thus, for example, duties to oneself such as developing one’s talents and not killing oneself can be deduced from the categorical imperative. So too can duties to others, such as keeping promises and helping those in need.

The first step in the categorical imperative procedure is for us to consider the underlying maxim of our proposed action. By *maxim*, Kant means the general rule in accordance with which the agent intends to act. For example, if I am thinking about assisting someone in need, my underlying maxim might be this: “When I see someone in need, I should assist him or her when it does not cause an undue burden on me.” The second step is to consider whether this maxim could be universalized to apply to everyone, such as “When anyone sees someone in need, that person should assist him or her when it does not cause an undue burden on the person.” If it can be universalized, then we accept the maxim, and the action is moral. If it cannot be universalized, then we reject the maxim, and the action is immoral. The general scheme of the CI procedure, then, is this:

Maxim of action



Universalize maxim



Accept successfully universalized maxim (reject unsuccessful maxim)

According to Kant, there is only one categorical imperative, but he presents three formulations of it:

- *Principle of the law of nature.* “Act as though the maxim of your action were by your will to become a universal law of nature.”
- *Principle of ends.* “So act as to treat humanity, whether in your own person or in that of any other, in every case as an end and never as merely a means.”
- *Principle of autonomy.* “So act that your will can regard itself at the same time as making universal law through its maxims.”

The theme that ties all of these formulations together is **universalizability**: Can a particular course of action be generalized so that it applies to any relevantly similar person in that kind of situation? For Kant, determining whether a maxim can successfully be universalized hinges on which of the three specific formulations of the categorical imperative that we follow. The bottom line for all three, though, is that we stand outside our personal maxims and estimate impartially and impersonally whether our maxims are suitable as principles for all of us to live by.

Let’s look at each of these formulations, beginning with the first and most influential, the principle of the law of nature.

The Principle of the Law of Nature: Four Examples

Again, the CI principle of the law of nature is this: “Act as though the maxim of your action were by your will to become a universal law of nature.” The emphasis here is that you must act analogous to the laws of physics, specifically insofar as such laws are not internally conflicting or self-defeating.⁵ For example, nature could not subsist with a law of gravity that had an object fall both up and down at the same time. Similarly, a system of morality could not subsist when a universalized maxim has an internal conflict. If you could consistently will that everyone would act on a given maxim, then there is an application of the categorical imperative showing the moral permissibility of the action. If you could not consistently will that everyone would act on the maxim, then that type of action is morally wrong; the maxim must then be rejected as self-defeated.

Kant gives four examples of the application of this test: (1) making a lying promise, (2) committing suicide, (3) neglecting one’s talent, and (4) refraining from helping others. The first and fourth of these are duties to others, whereas the second and third of these are duties to oneself. Kant illustrates how the CI principle of the law of nature works by applying it to each of these maxims.

Making a Lying Promise Suppose I need some money and am considering whether it would be moral to borrow the money from you and promise to repay it without ever intending to do so. Could I say to myself that everyone should make a false promise when he is in difficulty from which he otherwise cannot escape? The maxim of my act is M:

M. Whenever I need money, I should make a lying promise while borrowing the money.

Can I universalize the maxim of my act? By applying the universalizability test to M, we get P:

- P. Whenever anyone needs money, that person should make a lying promise while borrowing the money.

But, something has gone wrong, for if I universalize this principle of making promises without intending to keep them, I would be involved in a contradiction:

I immediately see that I could will the lie but not a universal law to lie. For with such a law [that is, with such a maxim universally acted on] there would be no promises at all. . . . Thus my maxim would necessarily destroy itself as soon as it was made a universal law.⁶

The resulting state of affairs would be self-defeating because no one in his or her right mind would take promises as promises unless there was the expectation of fulfillment. Thus, the maxim of the lying promise fails the universalizability criterion; hence, it is immoral. Now, I consider the opposite maxim, one based on keeping my promise:

- M1. Whenever I need money, I should make a sincere promise while borrowing it.

Can I successfully universalize this maxim?

- P1. Whenever anyone needs money, that person should make a sincere promise while borrowing it.

Yes, I can universalize M1 because there is nothing self-defeating or contradictory in this. So, it follows, making sincere promises is moral; we can make the maxim of promise keeping into a universal law.

Committing Suicide Some of Kant's illustrations do not fare as well as the duty to keep promises. For instance, he argues that the categorical imperative would prohibit suicide because we could not successfully universalize the maxim of such an act. If we try to universalize it, we obtain the principle, "Whenever it looks like one will experience more pain than pleasure, one ought to kill oneself," which, according to Kant, is a self-contradiction because it would go against the very principle of survival upon which it is based. But whatever the merit of the form of this argument, we could modify the principle to read "Whenever the pain or suffering of existence erodes the quality of life in such a way as to make nonexistence a preference to suffering existence, one is permitted to commit suicide." Why couldn't this (or something close to it) be universalized? It would cover the rare instances in which no hope is in sight for terminally ill patients or for victims of torture or deep depression, but it would not cover the kinds of suffering and depression most of us experience in the normal course of life. Kant seems unduly absolutist in his prohibition of suicide.

Neglecting One's Talent Kant's other two examples of the application of the CI principle of the law of nature are also questionable. In his third example, he claims that we cannot universalize a maxim to refrain from developing our talents. But again, could we not qualify this and stipulate that under certain circumstances it is permissible not to develop our talents? Perhaps Kant is correct in that, if everyone selfishly refrained from developing talents, society would soon degenerate into anarchy. But couldn't one universalize the following maxim M3?

M3. Whenever I am not inclined to develop a talent, and this refraining will not seriously undermine the social order, I may so refrain.

Refraining from Helping Others Kant's last example of the way the CI principle of the law of nature functions regards the situation of not coming to the aid of others whenever I am secure and independent. He claims that I cannot universalize this maxim because I never know whether I will need the help of others at some future time. Is Kant correct about this? Why could I not universalize a maxim never to set myself a goal whose achievement appears to require the cooperation of others? I would have to give up any goal as soon as I realized that cooperation with others was required. In what way is this contradictory or self-defeating? Perhaps it would be selfish and cruel to make this into a universal law, but there seems nothing contradictory or self-defeating in the principle itself. The problems with universalizing selfishness are the same ones we encountered in analyzing egoism, but it is dubious whether Kant's categorical imperative captures what is wrong with egoism. Perhaps he has other weapons that do elucidate what is wrong with egoism (we return to this later).

COUNTEREXAMPLES TO THE PRINCIPLE OF THE LAW OF NATURE

Kant thought that he could generate an entire moral law from his categorical imperative. The above test of universalizability advocated by Kant's principle of the law of nature seems to work with such principles as promise keeping and truth telling and a few other maxims, but it doesn't seem to give us all that Kant wanted. It has been objected that Kant's categorical imperative is both too wide and too unqualified. The charge that it is too wide is based on the perception that it seems to justify some actions that we might consider trivial or even immoral.

Counterexample 1: Mandating Trivial Actions

For an example of a trivial action that might be mandated by the categorical imperative, consider the following maxim M:

M. I should always tie my right shoe before my left shoe.

This generates the following principle P:

P. We should always tie our right shoe before our left shoe.

Can we universalize P without contradiction? It seems that we can. Just as we universalize that people should drive cars on the right side of the street rather than the left, we could make it a law that everyone should tie the right shoe before the left shoe. But it seems obvious that there would be no point to such a law—it would be trivial. But it is justified by the categorical imperative.

It may be objected that all this counterexample shows is that it may be *permissible* (not obligatory) to live by the principle of tying the right shoe before the left because we could also universalize the opposite maxim (tying the left before the right) without contradiction. That seems correct.

Counterexample 2: Endorsing Cheating

Another counterexample, offered by Fred Feldman,⁷ appears to show that the categorical imperative endorses cheating. Maxim M states:

M. Whenever I need a term paper for a course and don't feel like writing one, I will buy a term paper from Research Anonymous and submit it as my own work.

Now we universalize this maxim into a universal principle P:

P. Whenever anyone needs a term paper for a course and doesn't feel like writing one, the person will buy one from a suitable source and submit it as his or her own.

This procedure seems to be self-defeating. It would undermine the whole process of academic work because teachers wouldn't believe that research papers really represented the people who turned them in. Learning would not occur; grades and transcripts would be meaningless, and the entire institution of education would break down; the whole purpose of cheating would be defeated.

But suppose we made a slight adjustment to M and P, inventing M1 and P1:

M1. When I need a term paper for a course and don't feel like writing one, and no change in the system will occur if I submit a store-bought one, then I will buy a term paper and submit it as my own work.

P1. Whenever anyone needs a term paper for a course and doesn't feel like writing it, and no change in the system will occur if one submits a store-bought paper, then one will buy the term paper and submit it as one's own work.

Does P1 pass as a legitimate expression of the categorical imperative? It might seem to satisfy the conditions, but Kantian students have pointed out that for a principle to be universalizable, or lawlike, one must ensure that it is public.

However, if P1 were public and everyone was encouraged to live by it, then it would be exceedingly difficult to prevent an erosion of the system. Teachers would take precautions against it. Would cheaters have to announce themselves

publicly? In sum, the attempt to universalize even this qualified form of cheating would undermine the very institution that makes cheating possible. So, P1 may be a thinly veiled oxymoron: Do what will undermine the educational process in such a way that it doesn't undermine the educational process.

Counterexample 3: Prohibiting Permissible Actions

Another type of counterexample might be used to show that the categorical imperative refuses to allow us to do things that common sense permits. Suppose I need to flush the toilet, so I formulate my maxim M:

M. At time t_1 , I will flush the toilet.

I universalize this maxim:

P. At time t_1 , everyone should flush their toilet.

But I cannot will this if I realize that the pressure of millions of toilets flushing at the same time would destroy the nation's plumbing systems, and so I could not then flush the toilet.

The way out of this problem is to qualify the original maxim M to read M1:

M1. Whenever I need to flush the toilet and have no reason to believe that it will set off the impairment or destruction of the community's plumbing system, I may do so.

From this we can universalize to P1:

P1. Whenever anyone needs to flush the toilet and has no reason to believe that it will set off the destruction of the community's plumbing system, he or she may do so.

Thus, Kant seems to be able to respond to some of the objections to his theory.

Counterexample 4: Mandating Genocide

More serious is the fact that the categorical imperative appears to justify acts that we judge to be horrendously immoral. Suppose I hate people of a certain race, religion, or ethnic group. Suppose it is Americans that I hate and that I am not an American. My maxim is this:

M. Let me kill anyone who is American.

Universalizing M, we get P:

P. Always kill Americans.

Is there anything contradictory in this injunction? Could we make it into a universal law? Why not? Americans might not like it, but there is no logical contradiction involved in such a principle. Had I been an American when this command was in effect, I would not have been around to write this book, but the world

would have survived my loss without too much inconvenience. If I suddenly discover that I am an American, I would have to commit suicide. But as long as I am willing to be consistent, there doesn't seem to be anything wrong with my principle, so far as its being based on the categorical imperative is concerned.

As with the shoe-tying example, it would be possible to universalize the opposite—that no one should kill innocent people. Nevertheless, we certainly wouldn't want to say that it is permissible to adopt the principle "Always kill Americans."

We conclude, then, that even though the first version of the categorical imperative is an important criterion for evaluating moral principles, it still needs supplementation. In itself, it is purely formal and leaves out any understanding about the content or material aspect of morality. The categorical imperative, with its universalizability test, constitutes a necessary condition for being a valid moral principle, but it does not provide us with a sufficiency criterion. That is, if any principle is to count as rational or moral, it must be universalizable; it must apply to everyone and to every case that is relevantly similar. If I believe that it's wrong for others to cheat on exams, then unless I can find a reason to believe that I am relevantly different from these others, it is also wrong for me to cheat on exams. If premarital heterosexual coitus is prohibited for women, then it must also be prohibited for men (otherwise, with whom would the men have sex—other men's wives?). This formal consistency, however, does not tell us whether cheating itself is right or wrong or whether premarital sex is right or wrong. That decision has to do with the material content of morality, and we must use other considerations to help us decide about that.

OTHER FORMULATIONS OF THE CATEGORICAL IMPERATIVE

We've discussed Kant's first formulation of the categorical imperative; now we will consider the two others: the principle of ends and the principle of autonomy.

The Principle of Ends

Again, the principle of ends is this: "So act as to treat humanity, whether in your own person or in that of any other, in every case as an end and never as merely a means." Each person as a rational being has dignity and profound worth, which entails that he or she must never be exploited or manipulated or merely used as a means to our idea of what is for the general good (or to any other end).

What is Kant's argument for viewing rational beings as having ultimate value? It goes like this: In valuing anything, I endow it with value; it can have no value apart from someone's valuing it. As a valued object, it has *conditional* worth, which is derived from my valuation. On the other hand, the person who values the

object is the ultimate source of the object, and as such belongs to a different sphere of beings. We, as valuers, must conceive of ourselves as having *unconditioned* worth. We cannot think of our personhood as a mere thing because then we would have to judge it to be without any value except that given to it by the estimation of someone else. But then that person would be the source of value, and there is no reason to suppose that one person should have unconditional worth and not another who is relevantly similar. Therefore, we are not mere objects. We have unconditional worth and so must treat all such value-givers as valuable in themselves—as ends, not merely means. I leave it to you to evaluate the validity of this argument, but most of us do hold that there is something exceedingly valuable about human life.

Kant thought that this formulation, the principle of ends, was substantively identical with his first formulation of the categorical imperative, but most scholars disagree with him. It seems better to treat this principle as a supplement to the first, adding content to the purely formal CI principle of the law of nature. In this way, Kant would limit the kinds of maxims that could be universalized. Egoism and the principle regarding the killing of Americans would be ruled out at the very outset because they involve a violation of the dignity of rational persons. The process would be as follows:

1. Formulate the maxim (M).
2. Apply the ends test. (Does the maxim involve violating the dignity of rational beings?)
3. Apply the principle of the law of nature universalization test. (Can the maxim be universalized?)
4. Successful moral principles survive both tests.

In any event, we may ask whether the CI principle of ends fares better than the CI principle of the law of nature. Three problems soon emerge. The first has to do with Kant's setting such a high value on rationality. Why does reason and only reason have intrinsic worth? Who gives this value to rational beings, and how do we know that they have this value? What if we believe that reason has only instrumental value?

Kant's notion of the high inherent value of reason will be plausible to those who believe that humans are made in the image of God and who interpret that (as has the mainstream of the Judeo-Christian tradition) as entailing that our rational capabilities are the essence of being created in God's image: We have value because God created us with worth—that is, with reason. But, even nontheists may be persuaded that Kant is correct in seeing rationality as inherently good. It is one of the things rational beings value more than virtually anything else, and it is a necessary condition to whatever we judge to be a good life or an ideal life (a truly happy life).

Kant seems to be correct in valuing rationality. It does enable us to engage in deliberate and moral reasoning, and it lifts us above lower animals. Where he may have gone wrong is in neglecting other values or states of being that may have moral significance. For example, he believed that we have no obligations to

animals because they are not rational. But surely the utilitarians are correct when they insist that the fact that animals can suffer should constrain our behavior toward them: We ought not cause unnecessary harm. Perhaps Kantians can supplement their system to accommodate this objection.

This brings us to our second problem with Kant's formulation. If we agree that reason is an intrinsic value, then does it not follow that those who have more of this quality should be respected and honored more than those who have less?

- (1) Reason is an intrinsic good.
- (2) The more we have of an intrinsically good thing, the better.
- (3) Therefore, those who have more reason than others are intrinsically better.

Thus, by Kantian logic, people should be treated in exact proportion to their ability to reason, so geniuses and intellectuals should be given privileged status in society (as Plato and Aristotle might argue). Kant could deny the second premise and argue that rationality is a threshold quality, but the objector could come back and argue that there really are degrees in ability to use reason, ranging from gorillas and chimpanzees all the way to the upper limits of human genius. Should we treat gorillas and chimps as ends in themselves while still exploiting small babies and severely senile people because the former do not yet act rationally and the latter have lost what ability they had? If we accept the Kantian principle of ends, what should be our view on abortion, infanticide, and euthanasia?

Kant's principle of ends says all humans have dignity by virtue of their rationality, so they are permitted to exploit animals (who are intelligent but not rational). But suppose Galacticans who visited our planet were superrational, as superior to us as we are to other animals. Would we then be second-class citizens whom the Galacticans could justifiably exploit for their purposes? Suppose they thought we tasted good and were nutritious. Would morality permit them to eat us? Kantians would probably insist that minimal rationality gives one status—but then, wouldn't some animals who deliberate (chimps, bonobos, gorillas, and dolphins) gain status as persons? And don't sheep, dogs, cats, pigs, and cows exhibit minimally rational behavior? Should we eat them? (The Chinese think nothing is wrong with eating dogs and cats.)

There is a third problem with Kant's view of the dignity of rational beings. Even if we should respect them and treat them as ends, this does not tell us very much. It may tell us not to enslave them or not to act cruelly toward them without a good reason, but it doesn't tell us what to do in situations where our two or more moral duties conflict.

For example, what does it tell us to do about a terminally ill woman who wants us to help her die? What does it tell us to do in a war when we are about to aim our gun at an enemy soldier? What does it mean to treat such a rational being as an end? What does it tell us to do with regard to the innocent, potential victim and the gangsters who have just asked us the whereabouts of the victim? What does it tell us about whether we should steal from the pharmacy to procure medicine we can't afford in order to bring healing to a loved one?

It's hard to see how the notion of ends helps us much in these situations. In fairness to Kant, however, we must say that virtually every moral system has trouble with dilemmas and that it might be possible to supplement Kantianism to solve some of them.

The Principle of Autonomy

The final formulation of the categorical imperative is the principle of **autonomy**: "So act that your will can regard itself at the same time as making universal law through its maxims." That is, we do not need an external authority—be it God, the state, our culture, or anyone else—to determine the nature of the moral law. We can discover this for ourselves. And the Kantian faith proclaims, everyone who is ideally rational will legislate exactly the same universal moral principles.

The opposite of autonomy is **heteronomy**: The heteronomous person is one whose actions are motivated by the authority of others, whether it is religion, the state, his or her parents, or a peer group. The following illustration may serve as an example of the difference between these two states of being.

In the early 1960s, Stanley Milgram of Yale University conducted a series of social psychological experiments aimed at determining the degree to which the ordinary citizen was obedient to authority. Volunteers from all walks of life were recruited to participate in "a study of memory and learning." Two people at a time were taken into the laboratory. The experimenter explained that one was to play the role of the "teacher" and the other the role of the "learner." The teacher was put in a separate room from which he or she could see the learner through a window. The teacher was instructed to ask the learner to choose the correct correlate to a given word, and the learner was to choose from a set of options. If the learner got the correct word, they moved on to the next word. But, if the learner chose the wrong word, he or she was punished with an electric shock. The teacher was given a sample shock of 45 volts just to get the feeling of the game. Each time that the learner made a mistake, the shock was increased by 15 volts (starting at 15 volts and continuing to 450 volts). The meter was marked with verbal designations: slight shock, moderate shock, strong shock, very strong shock, intense shock, extreme-intensity shock, danger: severe shock, and XXX. As the experiment proceeded, the learner would generally be heard grunting at the 75-volt shock, crying out at 120 volts, begging for release at 150 volts, and screaming in agony at 270 volts. At around 300 volts, there was usually dead silence.

Now, unbeknown to the teacher, the learner was not actually experiencing any shocks; the learners were really trained actors simulating agony. The results of the experiment were astounding. Whereas Milgram and associates had expected that only a small proportion of citizens would comply with the instructions, 60 percent were completely obedient and carried out the experiment to the very end. Only a handful refused to participate in the experiment at all once they discovered what it involved. Some 35 percent left at various stages. Milgram's experiments were later replicated in Munich, Germany, where 85 percent of the subjects were found to be completely "obedient to authority."

There are two ways in which the problems of autonomy and heteronomy are illustrated by this example. In the first place, the experiment seems to show that the average citizen acts less autonomously than we might expect. People are basically heteronomous, herd followers. In the second place, there is the question about whether Milgram should have subjected people to these experiments. Was he violating their autonomy and treating them as means (rather than ends) in deceiving them in the way he did? Perhaps a utilitarian would have an easier time justifying these experiments than a Kantian.

In any case, for Kant, it is our ability to use reason in universalizing the maxims of our actions that sets rational beings apart from nonrational beings. As such, rational beings belong to a kingdom of ends. Kant thought that each of us—as a fully rational, autonomous legislator—would be able to reason through to exactly the same set of moral principles, the ideal moral law.

THE PROBLEM OF EXCEPTIONLESS RULES

One of the problems that plagues all formulations of Kant's categorical imperative is that it yields unqualified absolutes. The rules that the categorical imperative generates are universal and exceptionless. He illustrates this point with regard to truth telling: Suppose an innocent man, Mr. Y, comes to your door, begging for asylum, because a group of gangsters is hunting him down to kill him. You take the man in and hide him in your third-floor attic. Moments later the gangsters arrive and inquire after the innocent man: "Is Mr. Y in your house?" What should you do? Kant's advice is to tell them the truth: "Yes, he's in my house."⁸

What is Kant's reasoning here? It is simply that the moral law is exceptionless. It is your duty to obey its commands, not to reason about the likely consequences. You have done your duty: hidden an innocent man and told the truth when asked a straightforward question. You are absolved of any responsibility for the harm that comes to the innocent man. It's not your fault that there are gangsters in the world.

To many of us, this kind of absolutism seems counterintuitive. One way we might alter Kant here is simply to write in qualifications to the universal principles, changing the sweeping generalization "Never lie" to the more modest "Never lie, except to save an innocent person's life." The trouble with this way of solving the problem is that there seem to be no limits on the qualifications that would need to be attached to the original generalization—for example, "Never lie, except to save an innocent person's life (unless trying to save that person's life will undermine the entire social fabric)" or "Never lie, except to save an innocent person's life (unless this will undermine the social fabric)" or "Never lie, except to spare people great anguish (such as telling a cancer patient the truth about her condition)." And so on. The process seems infinite and time consuming and thus impractical.

However, another strategy is open for Kant—namely, following the *prima facie* duty approach advocated by twentieth-century moral philosopher William

D. Ross (1877–1971). Let's first look at the key features of Ross's theory and then adapt it to Kant's.

Ross and Prima Facie Duties

Today, Ross is perhaps the most important deontological theorist after Kant, and, like Pufendorf, Ross is a rule-intuitionist. There are three components of Ross's theory. The first of these is his notion of "moral intuition," internal perceptions that both discover the correct moral principles and applies them correctly. Although they cannot be proved, the moral principles are *self-evident* to any normal person upon reflection. Ross wrote,

That an act, qua fulfilling a promise, or qua effecting a just distribution of good . . . is prima facie right, is self-evident; not in the sense that it is evident . . . as soon as we attend to the proposition for the first time, but in the sense that when we have reached sufficient mental maturity and have given sufficient attention to the proposition it is evident without any need of proof, or of evidence beyond itself. It is evident just as a mathematical axiom, or the validity of a form of inference, is evident. . . . In our confidence that these propositions are true there is involved the same confidence in our reason that is involved in our confidence in mathematics. . . . In both cases we are dealing with propositions that cannot be proved, but that just as certainly need no proof.⁹

Just as some people are better perceivers than others, so the moral intuitions of more reflective people count for more in evaluating our moral judgments. "The moral convictions of thoughtful and well-educated people are the data of ethics, just as sense-perceptions are the data of a natural science."¹⁰

The second component of his theory is that our intuitive duties constitute a plural set that cannot be unified under a single overarching principle (such as Kant's categorical imperative or the utilitarian highest principle of "the greatest good for the greatest number"). As such, Ross echoes the intuitionism of Pufendorf by presenting a list of several duties, specifically these seven:

1. Promise keeping
2. Fidelity
3. Gratitude for favors
4. Beneficence
5. Justice
6. Self-improvement
7. Nonmaleficence

The third component of Ross's theory is that our intuitive duties are not absolute; every principle can be overridden by another in a particular situation. He makes this point with the distinction between **prima facie duties** and **actual duties**. The term *prima facie* is Latin for "at first glance," and Ross's point is that all

seven of the above listed moral duties are tentatively binding on us until one duty conflicts with another. When that happens, the weaker one disappears, and the stronger one emerges as our actual duty. Thus, although *prima facie* duties are not actual duties, they may become such, depending on the circumstances. For example, if we make a promise, we put ourselves in a situation in which the duty to keep promises is a moral consideration. It has presumptive force, and if no conflicting *prima facie* duty is relevant, then the duty to keep our promises automatically becomes an actual duty.

What, for Ross, happens when two duties conflict? For an absolutist, an adequate moral system can never produce moral conflict, nor can a basic moral principle be overridden by another moral principle. But Ross is no absolutist. He allows for the overridability of principles. For example, suppose you have promised your friend that you will help her with her homework at 3:00 p.m. While you are on your way to meet her, you encounter a lost, crying child. There is no one else around to help the little boy, so you help him find his way home. But, in doing so, you miss your appointment. Have you done the morally right thing? Have you broken your promise?

It is possible to construe this situation as constituting a conflict between two moral principles:

1. We ought always to keep our promises.
2. We ought always to help people in need when it is not unreasonably inconvenient to do so.

In helping the child get home, you have decided that the second principle overrides the first. This does not mean that the first is not a valid principle—only that the “ought” in it is not an absolute “ought.” The principle has objective validity, but it is not always decisive, depending on which other principles may apply to the situation. Although some duties are weightier than others—for example, non-maleficence “is apprehended as a duty of a more stringent character . . . than beneficence”—the intuition must decide each situation on its own merits.

Kant and the *Prima Facie* Solution

Many moral philosophers—egoists, utilitarians, and deontologists—have adopted the *prima facie* component of Ross’s theory as a convenient way of resolving moral dilemmas. In doing so, they typically don’t adopt Ross’s account of moral intuitions or his specific set of seven duties (that is, the first two components of Ross’s theory). Rather, they just incorporate Ross’s concepts of *prima facie* duty and actual duty as a mechanism for explaining how one duty might override another.

How might this approach work with Kant? Consider again Kant’s innocent-man example. First, we have the principle L: “Never lie.” Next, we ask whether any other principle is relevant in this situation and discover that that is principle P: “Always protect innocent life.” But we cannot obey both L and P (we assume for the moment that silence will be a giveaway). We have two general principles; neither of them is to be seen as absolute or nonoverridable but rather as *prima*

facie. We have to decide which of the two overrides the other, which has greater moral force. This is left up to our considered judgment (or the considered judgment of the reflective moral community). Presumably, we will opt for P over L, meaning that lying to the gangsters becomes our actual duty.

Will this maneuver save the Kantian system? Well, it changes it in a way that Kant might not have liked, but it seems to make sense: It transforms Kant's absolutism into a *modest objectivist* system (as described in Chapter 3). But now we need to have a separate criterion to resolve the conflict between two competing *prima facie* principles. For Ross, moral intuitions performed that function. Since Kant is more of a *rational* intuitionist, it would be the job of reason to perform that function. Perhaps his second formulation of the categorical imperative—the principle of ends—might be of service here. For example, in the illustration of the inquiring killer, the agent is caught between two compelling *prima facie* duties: “Never lie” and “Always protect innocent life.” When determining his actual duty, the agent might reflect on which of these two duties best promotes the treatment of people as ends—that is, beings with intrinsic value. This now becomes a contest between the dignity of the would-be killer who deserves to hear the truth and the dignity of the would-be victim who deserves to live. In this case, the dignity of the would-be victim is the more compelling value, and the agent's actual duty would be to always protect innocent life. Thus, the agent should lie to protect the life of the would-be victim.

THE PROBLEM OF POSTERITY

In the chapters on egoism (Chapter 6) and utilitarianism (Chapter 7), we explored the problem of posterity: What obligations do we owe to future generations? This question is especially relevant now as we wrestle with issues of environmental responsibility. People today can undermine the environment in ways that may not seriously impact us while we're alive but will have devastating consequences on the well-being of future generations. The issue concerns the moral obligations that we have toward people who do not yet exist. We've seen that egoists, with their emphasis on personal self-interest, gain nothing by preserving natural resources for future generations that do not yet exist and thus can give no benefit to the egoist. Utilitarians, we've seen, do better: We have a duty to maximize general happiness, and it does not matter that we cannot identify these future people.

What, though, do deontologists say about the problem of posterity? Intuitionists such as Butler or Ross might find it self-evident that we do or we don't have obligations to future generations; thus, they are not much help in offering a systematic account of our duty to posterity. The situation is not much better with Kant who, with his strong emphasis on particular rational people, would have a particularly difficult time generating principles that would require duties to future agents. Kant seems to require identifiable people as the objects of our duties.

Here, though, is one solution that might be extracted from Kant's moral philosophy. Like earlier deontologists of his time such as Pufendorf, Kant

distinguishes between perfect and imperfect duties. *Perfect duties* are those that require precise behavior toward particular people. For example, my duty to tell the truth is a perfect duty because it mandates in each situation that I should be truthful to each person who I communicate with. *Imperfect duties*, by contrast, do not require precise behavior toward particular people. An example of this would be my duty to be charitable. Although I should help those in need, I am not obligated to help *everyone* in need or even any *particular person* in need. It is up to me to determine who the receivers of my charity should be.

Applying this distinction to the problem of posterity, we might recognize an imperfect duty to promote the well-being of people who will exist in the future—even if we don't know who they are. As long as we are reasonably sure that they will exist, we have a duty to see to it that their lives are reasonably good. Although this helps somewhat in addressing the problem of posterity, it is not the best solution. Because imperfect duties are so undefined, they may be overridden by perfect duties. Thus, my imperfect duty to preserve the environment for future generations may be overridden by my perfect duty of justice to allow existing people in developing countries to exploit the environment for their economic advancement. This dilemma regarding developing countries is in fact at the heart of much of the environmental debate today.

Here's another solution offered by some Kantians: We do have obligations to the future world in the persons of our children. We have an obligation to leave the environment in good shape for our children, and they in turn will have such an obligation for their children, and so on, so that the question of posterity is taken care of. The duty carries over to future generations. But this seems to commit a fallacy of transitivity:

If A has a duty X to B and if B has a duty X to C, then A has a duty X to C.

This formula is invalid. I may have a duty to keep my promise to you, and you have a duty to keep your promise to your neighbor, but I don't have a duty to keep my promise to you to your neighbor. Similarly, our generation may have a duty to provide the next generation (our children) with an adequate living, but we don't have an obligation to provide the next generation with one. In fact, given scarce resources, it may be that simply sharing our resources with the next generation will use up a considerable amount so that the next generation will be unable to pass down a sufficient amount to the next generation. Considering the importance of the problem of posterity in environmental ethics today, defenders of Kant will undoubtedly wish to explore other more successful solutions.

CONCLUSION: A RECONCILIATION PROJECT

Utilitarianism and deontological systems such as Kant's are radically different types of moral theories. Some people seem to gravitate to the one and some to the other, but many people find themselves dissatisfied with both positions. Although they see something valid in each type of theory, at the same time there

is something deeply troubling about each. Utilitarianism seems to catch the spirit of the purpose of morality, such as human flourishing and the reduction of suffering, but undercuts justice in a way that is counterintuitive. Deontological systems seem right in their emphasis on the importance of rules and the principle of justice but tend to become rigid or to lose focus on the central purposes of morality.

One philosopher, William Frankena, has attempted to reduce this tension by reconciling the two types of theories in an interesting way. He calls his position “mixed deontological ethics” because it is basically rule centered but in such a way as to take account of the teleological aspect of utilitarianism.¹¹ Utilitarians are right about the purpose of morality: All moral action involves doing good or alleviating evil. However, utilitarians are wrong to think that they can measure these amounts or that they are always obligated to bring about the “greatest balance of good over evil,” as articulated by the principle of utility.

In place of the principle of utility, Frankena puts forth a near relative, the *principle of beneficence*, that calls on us to strive to do good without demanding that we be able to measure or weigh good and evil. Under his principle of beneficence, he lists four hierarchically arranged subprinciples:

1. One ought not to inflict evil or harm.
2. One ought to prevent evil or harm.
3. One ought to remove evil.
4. One ought to do or promote good.

In some sense, subprinciple 1 takes precedence over 2, 2 over 3, and 3 over 4, other things being equal.

The *principle of justice* is the second principle in Frankena’s system. It involves treating every person with equal respect because that is what each is due. To quote John Rawls, “Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. . . . The rights secured by justice are not subject to political bargaining or to the calculus of social interests.”¹² There is always a presumption of equal treatment unless a strong case can be made for overriding this principle. So even though both the principle of beneficence and the principle of justice are *prima facie* principles, the principle of justice enjoys a certain priority. All other duties can be derived from these two fundamental principles.

Of course, the problem with this kind of two-principle system is that we have no clear method for deciding between them in cases of moral conflict. In such cases, Frankena opts for an intuitionist approach similar to Ross’s: We need to use our intuition whenever the two rules conflict in such a way as to leave us undecided on whether beneficence should override justice. Perhaps we cannot decisively solve every moral problem, but we can solve most of our problems successfully and make progress toward refining our subprinciples in a way that will allow us to reduce progressively the undecidable areas. At least, we have improved on strict deontological ethics by outlining a system that takes into account our intuitions in deciding complex moral issues.

CASE STUDY 13: Who shall live?

Here is a dilemma facing the medical world today: establishing priorities amid a scarcity of resources, in this case a scarcity of kidney machines in a particular hospital. Some years ago, this scarcity resulted in four people applying (without realising this, of course) for what was at the time the sole available dialysis machine in the renal unit. Obviously, only one of the four could be admitted. The others might be able to have their houses made suitable for the installation of a home dialysis machine, and the possibility of a kidney transplant existed for most, if not all, of them: but both these procedures would require time, during which any one of them could die. The only way to be sure that this did not occur for one of them was to put him or her *immediately* (and regularly thereafter) on the available machine.

How would you set about making the decision? Imagine that you are one of the hospital team who has to decide between the four sick people. The script which follows contains the verbatim reports on all four written by the medical social worker. Only the names have been changed. Also available to the team will be the patient's medical records. These will show whether any of the four has any other serious ailment: if so, that patient would be excluded from the dialysis machine because, harsh though it may sound, it would be unfair to any of those refused access, and to their relatives, if he/she were to die soon afterwards of renal failure, while the person selected died at the same time, not of renal failure, but of something else such as heart or lung failure. We can take it, since these records are naturally not available to us, that there was nothing else seriously wrong with the patients. Besides the medical social worker, the team would be likely to include the surgeons, registrar, doctor(s), sister, senior nurses, and any others working in or attached to the renal unit, such as the home dialysis administrator, whose job is to assess the suitability of patients' homes for the installation of a machine.

So your task is straightforward: who shall live – or, put less dramatically, which of the four will you choose to ensure does not die of renal failure because of lack of access to a dialysis machine? More importantly, on what basis, by the application of what criteria, do you reach your decision? How far do you consider the categorical imperative to be relevant and applicable? Can you establish a universal law or principle that may ease the burden of those involved in the making of such grim, but humanly necessary (as matters are), decisions?

The most rewarding use of this study would be to do it with a group of friends or colleagues. You might proceed as follows:

- 1 Read the records carefully.
- 2 Without any discussion whatsoever, take a straw poll on which person you would choose.
- 3 Discuss the grounds on which each person's choice was made, and see if

you can agree on any kind of priority of criteria (age, family responsibility and support, usefulness to the community, etc.)

- 4 Take a final vote, and see if there has been a switch since the second stage. You may find it necessary to take a series of votes, each time eliminating the last on the list, in order to reach a decision with the maximum amount of agreement possible.

Case No. 1	John White
Age	26 years
Employment	Part-time university lecturer, part-time researcher working for PhD in psychology.
Family circumstances	Family immigrated when John, only child, was of school age. Father now dead, mother has returned to Barbados.
Housing	Rented bedsitter in cosmopolitan street. Poor amenities, stove on landing, etc.; lavatory shared with eight other tenants; no bathroom. After another year, Mr White will become entitled to accommodation within the university.

Medical social worker's report

I visited Mr White in his bedsitter. Although he is obviously a very fastidious young man and has done his best to make his room habitable, the condition of the house is so bad that his efforts are largely wasted. Apparently, the landlord owns several houses in the street and lets rooms to tenants, his sole means of income. He has repeatedly refused to do anything to repair the state of these houses. The roof leaks, the plumbing is inefficient, and there is dry rot on the stairs. Despite his depressing surroundings, Mr White is a cheerful young man. It appears that he considers himself lucky to have found anywhere to live at all. I spoke to him about his mother; due to homesickness she returned home after his father's death three years ago. Mr White sends her part of his salary each month. Mr White has adjusted to the fact of his serious illness very quickly. Through his contacts at the university he has gained a good deal of information about the routine use of a kidney machine and assured me that he would be able to adjust his working hours to fit in with the hospital schedule. As he is young, he may be suitable for a transplant operation later on. One of the problems is his lack of a permanent home. Judging by his general attitude towards his tenants, the landlord would hardly tolerate the alterations that would be required for a home kidney machine to be installed – in any case, it would mean that the whole house would have to be rewired and have new plumbing. Unfortunately, Mr White is not in a financial position to gain a mortgage in order to buy his own property, though

the university may help here – I gather that he is highly thought of there and, until his illness, was expected to achieve a senior post within a fairly short time.

Signed – Amy Collins

Case No. 2	William Johnston
Age	52 years
Employment	Crane driver, shift work 6 am–2 pm, 2 pm–10 pm alternate weeks.
Family circumstances	Wife, 49, works as cleaner, mornings only. Four children; two married daughters, one lives near the family home; two sons, 13 and 15 years – the elder plans to leave school in the summer, the younger is academically inclined, has prospects of going to college. Also grandmother (Mr Johnston’s mother), 83, arthritic, recently had cataract operation.
Housing	Rented house in decaying terrace. Very shabby outside, plaster falling off walls, etc. Family have been on housing list for fifteen years.

Medical social worker’s report

Mrs Johnston is a plump, homely woman, devoted to her family. The house shows signs of do-it-yourself activities, for example, Mr Johnston laid new boards on the living-room floor, and, at the time when he was taken ill, was about to redecorate the kitchen. Mrs Johnston told me that he has to do this at least once a year as damp comes through the outside wall continuously – she showed me a corner where mould was forming on the paintwork. The boys were at school when I called, but I got the impression that this is a happy family. Mrs Johnston is a placid personality and shows a wonderful sense of humour. Mr Johnston’s mother can be very difficult at times but I got the impression that Mrs Johnston does not let this get her down. She reacted vehemently against my asking whether she had considered placing the old lady in a home, so that she would not find the additional strain of tending her sick husband so great. Money is a problem at the moment; Mr Johnston’s firm are not supplementing his sick pay and the weekly income is reduced by a third. But Mrs Johnston is more worried about what job her husband will be fit to do when he comes out of hospital than about money itself; she says that her husband will never settle for early retirement and she cannot see the firm giving him a less physically demanding job. I promised to look into this. She also enquired about the possibility of a transplant operation. I told her as gently as I could that her husband is rather too old to be considered for this, though of course he is basically a strong man.

Signed – Amy Collins

Case No. 3	Sue Gibson
Age	37 years
Employment	Civil Servant; senior position in department of economic statistics.
Family circumstances	Unmarried. Lives with and supports aged mother (79) and sister with Down's Syndrome (30).
Housing	A rather dark two-bedroom flat; the sisters share a bedroom. Flat owned by Miss Gibson and her mother. Apart from a married brother who lives and works abroad, there are no other relatives.

Medical social worker's report

I visited Mrs Gibson and her younger daughter quite early in the morning. Mrs Gibson is obviously at a loss without Sue; the Down's Syndrome daughter is difficult to handle and is too strong for Mrs Gibson to control. I witnessed a struggle between the two as Mrs Gibson tried to persuade her daughter to dress. Mrs Gibson is a thin, wiry little woman, very shy by nature, she whispered answers to my questions. I think she feels that she has had an unlucky deal in life. Her husband was killed when the children were quite young, the younger daughter has always been a problem and would have been placed in an institution long ago had not Sue rebelled against the idea. Now that Sue too has ill health Mrs Gibson feels that her world is falling about her. Sue has been a strong emotional prop as well as the breadwinner for the last twenty years. I spoke to Mrs Gibson about the feasibility of installing a home kidney unit – this would mean Sue having a room to herself and a water tank with a larger capacity would be an essential. Mrs Gibson said that she would be prepared to share her bedroom with her younger daughter but I sensed that she felt misgivings on the subject – after all, she is now an old lady and finds her younger daughter's noisy ways very wearing. It seems to me that the only hope is for Sue to be persuaded that her sister should be removed from the family home – on the grounds of her mother's health if not her own.

Signed – Amy Collins

Case No. 4	Jim Spencer
Age	17 years
Employment	Assistant baker 6 am–2 pm
Family circumstances	Lives with parents. Father (47) is an insurance agent. Mother used to work part-time in a children's nursery but gave up due to nervous troubles – she is still receiving outpatient treatment. There was a daughter, two years younger than Jim, she was killed eighteen months ago in a road accident.

in his life (except, possibly, his hypochondria) was allowed to interfere with this. Most of us, however, are committed to a variety of activities and social groups, all of which make demands on us from the duty angle. Kant in his writings seems to suggest that if we face a clash of duties we shall know what takes precedence, simply by following the categorical imperative (pp. 104ff). But does this really work? Talk (or think) this through with some of the following dilemmas in mind:

- (a) It is your turn to give an oral presentation at one of your seminars, and you know you will only just make it in time. On the way you witness a road accident, and are asked to stand by and make a statement.
- (b) All the people in your office fiddle their expenses, and for one not to do the same would mean exposing this practice to those in higher authority. Yet your conscience tells you not to involve yourself in this activity.
- (c) You have some important work to catch up on urgently, but a close friend, whom you now see only rarely, since she moved away, turns up unexpectedly with a couple of hours to spare on her way elsewhere.
- (d) 'If I had to choose between betraying my country and betraying my friend I hope I should have the guts to betray my country.' – E. M. Forster. Do you agree?
- (e) A married couple wish to separate, but have children at an impressionable age.
- (f) You have run up debts which you can't repay without sacrificing your student career and returning to a boring, non-demanding job. (In days gone by, men in this predicament were expected to commit suicide, but you may think that to be taking the idea of duty somewhat too seriously.)
- (g) You are a nurse who believes strongly in the sanctity of life. You are instructed to give 'nursing care only' to a newborn, profoundly handicapped infant (meaning that you are to allow it to die as comfortably as possible).
- (h) Similarly, you are told by relatives of a dying patient not to tell him that his illness is terminal: but he constantly asks you for his prognosis.
- (i) You share a sexual attraction with someone other than your partner, whom you are sure will not find out if you are unfaithful.

CASE STUDY 15: The case of the lying chancellor

In September, 1949, the pound sterling was devalued against the US dollar. In those days, the pound did not 'float' or 'find its market level' as it does today: the rate of exchange was fixed, and could be altered (and this happened

rarely) only by the decision of the chancellor of the exchequer. In 1949 the chancellor was Sir Stafford Cripps, one of the towering figures of the Labour Party, and among the most brilliant intellectuals ever to have entered the House of Commons. He was in addition a deeply religious man, austere, puritanical, even remote. (Winston Churchill is said to have remarked biting about him: 'There, but for the Grace of God, goes God.')

For weeks before the devaluation was announced there had been speculation that this was due to happen. Consistently throughout this period Cripps denied that there were any such plans. To press and public alike his reply was invariably the same: 'There will be no devaluation.' He maintained this stance right up to the moment when he went to the Despatch Box in the House of Commons and made his historic announcement. The debate to which this led was one of the most memorable in the history of Parliament, but the political aspects of the issue are not our concern. (If you are interested you can read a full account of the event in Michael Foot's biography of Aneurin Bevin, Vol. II, pp. 266–73.)

For many people at the time (and there are plenty who still affirm this) the main cause of sorrow at the announcement was not the fall in value of the pound in their pockets but the devaluation of a respected man's reputation. Cripps was known internationally as a man of honour and integrity; he was accepted generally as being in the George Washington league of those who 'could not tell a lie'. Yet with one brief speech he revealed that for weeks he had been lying openly, deliberately, and consistently; and not just to his friends and colleagues but to the nation and the world. It was a bitter blow for any idealists who held that such matters as truth-telling should (and therefore could: it was Kant who is said to have remarked 'Ought implies can') always occur – in fact, be taken for granted – in every feature of human intercourse. To be let down by one of the most respected standard bearers of this approach made the blow so much more deadly. It was as if Immanuel Kant had begun not only to throw refuse into his neighbours' gardens but had then accused them of stealing it.

By others at the time, however, this reaction was looked upon as totally unrealistic, taking no account of the desperate dilemma that faced the chancellor. What, it was asked, could Cripps realistically have done instead? If he had said, in reply to press questions, 'Yes, on such-and-such a date I shall be devaluing the pound by x per cent', everyone in the country would have been exchanging their pounds for dollars and the economic consequences for the country would have been disastrous. If he had said 'No comment', people would have jumped to their own conclusions, with the same consequences. If he had said 'I don't know', he would still, at least in the later stages, have been telling a lie, nobody would have believed him, and speculation of almost the same proportions would have occurred. He could have resigned rather than lie, but again people would have drawn their own conclusions from such an act; after all, there had been talk of devaluation throughout the whole summer,

and a resignation of that magnitude would not have been interpreted at that time as a purely accidental coincidence. The choice, therefore, before the chancellor was either economic disaster for the nation, or a lie on his conscience. The only way out of this dilemma was not to have been there in the first place!

What is your opinion? Do you think that sometimes people have to compromise with even the most high-sounding and universally accepted principles? If so, on what basis would you make the decision that such a compromise is right or desirable in one set of circumstances, but not in another? If not – that is, if you believe that people should tell the truth in all circumstances, whatever the consequences – have you thought what this might imply concerning people's daily communication with each other? Supposing an old lady of whom you are very fond bakes you a cake which you find excruciating and then, desperately anxious to have pleased you, she asks you later whether you enjoyed it: what do you say? Is a doctor acting wrongly if (s)he untruthfully answers a dying patient's question about his prospects, out of fear that the patient could not cope mentally with the truth? If your partner buys an expensive new garment which you dislike intensely, do you tell him/her the truth – the whole truth – when asked your opinion of it?

See if you can universalise truth-telling in such terms that nobody should be caused unnecessary pain: something like 'Always tell the truth unless by so doing you will needlessly hurt another person.' If this were a categorical imperative, would you ever be sure that others were not lying in order to spare your feelings? Could you ever believe anything that anybody ever told you about yourself unless this caused you pain? (And how do you qualify, or quantify, 'needlessly' in this context?)

Can you express any categorical imperative other than in the form of 'Do not do x except when . . .'? ('Do not steal except when this will save a life', etc.) How categorical is the resultant so-called imperative? Is Kant logically wrong?