

# UTILITARIAN ETHICS

---

---

“ASK YOURSELF WHETHER YOU ARE HAPPY,  
AND YOU CEASE TO BE SO.”

– John Stuart Mill (1806-1873)

## THE GREATEST HAPPINESS PRINCIPLE

Suppose you're visiting a friend at his cabin retreat up in the mountains. It's January, with lots of snow on the ground, and you have an hour to yourself while your friend buys supplies in town some twenty miles down the valley. The snow is blowing hard enough to keep you inside, and there's not much to do there – no TV, the hot tub's out of order, and all the books are written in either Greek or Sanskrit. But you do notice a couple of CD's lying by the CD player. You walk over to glance at the titles: one is a Barry Manilow album, the other is Shostakovich's 3rd Symphony. Barry Manilow, you reflect, is one of your favorite recording artists – indeed, some of your fondest memories involve listening to his unforgettable “Mandy.” Shostakovich, on the other hand, you despise as a noisy and disoriented Russian. You would rather poke out your eyes with a sharp stick than listen to an hour of Shostakovich. So: Which CD should you play?

Is there even a question to be asked here? Isn't it obvious that you ought to listen to Barry Manilow? That is what will give you the most pleasure, and so that's the CD you should play. And if in the end you decide to play the Shostakovich instead, it will be in pursuit of yet some other pleasure: Perhaps you're impressed that your friend – whose musical taste you deeply respect – has this particular CD; or you might want to be able to discuss the CD when your friend returns; and so on. But in each of these cases, it is apparent that you will do whatever you think will maximize your happiness (at least in the short term, preferably in the long term).

Utilitarianism is no different from this kind of reasoning, except that it adds **impartiality**, claiming that my happiness is of no greater or lesser importance than the happiness of anyone else. So utilitarianism, we might say as a first approximation, is no different from simple prudential reasoning, altered by the impartiality principle.<sup>1</sup>

### Allocating Scarce Resources

A common example of utilitarianism in action is where some third party needs to allocate scarce resources among a group of individuals, none of whom have a special claim to that resource. How should it be divided? The rule nearly always followed is the utilitarian principle of maximizing the overall happiness. Take kidneys, for example. Most of us are born with two, but sometimes we need a replacement, and life on a hemodialysis machine lasts only so long. Since the first kidney transplant in 1954, over 100,000 have been performed, with a current success rate of 93 percent (in comparison, there have been about 6,000 liver transplants with a success rate of 75 percent, and 8,000 heart transplants with a success rate of 82 percent).

Kidneys for transplantation come either from the recently deceased or from living donors. You need only one kidney to lead a healthy life, and so people are allowed to give (and in some countries, to sell) one of their kidneys. Even still, there are not nearly enough kidneys to meet the demand. At any given time in the United States, 36,000 people are in need of a kidney, with only 10,000 available.<sup>2</sup> Given the limited resources, how does the medical community (or we as the larger society)

---

<sup>1</sup> Impartiality comes in many varieties. I might adopt an attitude of impartiality towards myself and my best friend, or towards all of my friends (while privileging myself above them all), or towards myself and my family, or all my neighbors, and so on. Utilitarianism requires that this impartiality extend to all sentient creation.

<sup>2</sup> This is not for lack of kidneys, but of willing donors. Less than one percent of those who die in the United States donate their organs. While many of these aren't suitable organ donors, of the roughly 23,000 who die each year from brain death (and thus typically have healthy organs to donate), only 4,000 donate their organs.

decide how to parcel out those kidneys? This allocation effectively decides who lives and who dies. Should they be distributed on a “first come, first serve” basis? (But what if the person at the top of the list is 95 years old with a failing heart and Alzheimer’s, and the person at the bottom of the list is an otherwise healthy five year old?) Does it matter if the recipient has to care for dependents at home? Does it matter if the recipient is popular or well liked in the community? Devising a good allocation scheme is far from easy, but one principle that nearly always finds its way into such a scheme is the utilitarian principle: Everything else being equal, allocate the kidneys in such a way as to maximize overall well-being. Put them where they will bring about the most good. This is a principle that makes good sense to many people.

## UTILITARIANISM AS EMPIRICAL

**Jeremy Bentham** viewed moral theory as an empirical project: simply look about you, and see what it is that human beings find good and bad. What you’ll discover, Bentham claims, is that humans desire pleasure and abhor pain. As he wrote in his *Introduction to the Principles of Morals and Legislation* (1789):

Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do.

We evaluate actions in terms of the amount of pain or pleasure that they produce, and each pain and pleasure is evaluated in terms of the following criteria: “its intensity, its duration, its certainty or uncertainty, its propinquity or remoteness, its fecundity, its purity, and its extent — that is, the number of persons ... who are affected by it.” In order to keep these criteria well in mind, the ever-resourceful Bentham devised the following mnemonic poem for social reformers everywhere:

*Intense, long, certain, speedy, fruitful, pure* —  
Such marks in *pleasures* and in *pains* endure.  
Such pleasures seek if *private* be thy end:  
If it be *public*, wide let them *extend*.  
Such *pains* avoid, whichever be thy view:  
If pains *must* come, let them *extend* to few.

In more general terms, Bentham explained utilitarianism as follows:

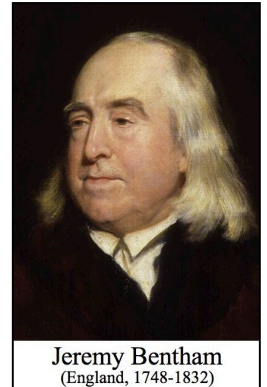
By the Principle of Utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question; or what is the same thing in other words, to promote or to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

## THE RIGHT AND THE GOOD

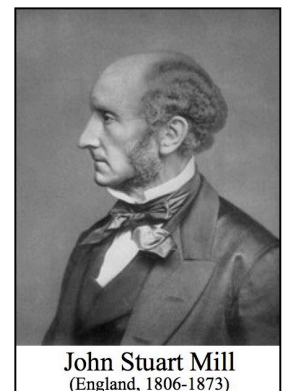
Utilitarianism is a **consequentialist** moral theory in that it decides the moral worth of an action solely on the basis of its consequences. Motives, intentions, the character of the agent — none of this ultimately matters in morally evaluating an action. A world filled with virtuous people acting always with good intentions will likely be a better world (insofar as it contains more of what is good); but such virtue and such intentions are worthy or desirable only so far as they increase this good — according to utilitarianism.

The general goal of consequentialist theories is to maximize the good; but what is this good that we are to maximize? Bentham believed that the good was pleasure, and in this regard he was closely followed by **John Stuart Mill** (1806-1873), whose father was a good friend of Bentham’s, and who became the leading advocate of utilitarianism in the generation following Bentham. In Mill’s short work entitled *Utilitarianism* (1861) we find one of the clearest and ablest discussions of utilitarianism, and it is this text that we will be considering in what follows.

Mill based utilitarianism on what he called the **greatest happiness principle (GHP)**: the right action among the alternatives open to us is the action that results in the greatest happiness for the greatest number of people. This happiness was good



Jeremy Bentham  
(England, 1748-1832)



John Stuart Mill  
(England, 1806-1873)

in and of itself, something desired by all, and therefore our **final good**; every other good is only an **instrumental good**, as a means to happiness.

Both Mill and Bentham often write about maximizing happiness, but what they mean by happiness is always something very definite, namely, **pleasure and the absence of pain**. So utilitarianism is a form of **hedonism** (from the Greek word *hêdonê*, which means “pleasure”). The good is pleasure, which ultimately is just a certain kind of psychological state. Our actions will cause various people to experience pleasures and pains, and morally right actions are simply whichever will maximize pleasures and minimize pains, however that is managed.

Other utilitarians have wished to remain neutral as to what it is that humans actually want, and so they view the good simply as the satisfaction of one’s preferences: thus, an action is right insofar as it satisfies as many preferences as possible. These are the two most common conceptions of the good among utilitarians: pleasure and preference-satisfaction. In reading Mill, one can detect both of these understandings of the good. In particular, when Mill distinguishes between different kinds of pleasure, he seems to be basing his argument on the satisfaction of different preferences (see below).

### MAXIMIZING THE NET GOOD, OVER THE LONG HAUL, FOR EVERYONE

A few possible ambiguities surrounding utilitarianism should be addressed immediately. One concerns “maximizing the good.” Presumably we want to maximize not the *total* good, but rather the balance of good over bad, or the *net* good. For instance, in running a business, the goal is not to maximize the total income but rather to maximize the profit (the net income). It’s obviously better to have \$500,000 in income and \$100,000 in expenses (for a net profit of \$400,000) than to have \$1 million in income and \$1 million in expenses (with no net profit).

This sort of consequentialism is intuitively plausible as a moral foundation. Consider Leibniz’s God creating the best of all possible worlds: presumably such a world will have the greatest balance of good over evil, since God is the source of the good, and would not allow more evil than necessary. Insofar as we want to do the right thing, it seems that we would want to emulate such God-like behavior and strive to increase the good and lessen the evil in the world.

Utilitarians also have the long view in mind when they speak of maximizing the net good. Actions that bring about a great deal of pleasure in the short run but which lead to considerable misery in the long run (say, addicting yourself to heroin) are not endorsed by utilitarians. Just how long a view the utilitarian should take, of course, is a question needing discussion (we will come back to this when we consider the problem of calculating the likely consequences of our actions).

Finally, utilitarianism does not place any special weight on the pleasures and pains of the agent. The greatest happiness principle refers not to the happiness of any one individual, but rather to the happiness of all humans — and, where possible, “to the whole sentient creation” (that is, to all creatures with the ability to experience pleasure and pain). This incorporates the “impartiality” criterion that is central to most ethical systems. Here, what we treat impartially are the pleasures and pains of each individual. Because Mill’s utilitarianism views the good as pleasure, we classify it as a kind of **hedonism**. But there are two broad kinds of hedonism: *private (egoistic)* and *social*. The first merely enjoins us to maximize our own pleasure while the second enjoins us to maximize the sum-total of everyone’s pleasure. The impartiality of Mill’s hedonism marks it as social.

### ACT VS RULE UTILITARIANISM

More recent moral theorists have distinguished between two different kinds of utilitarianism: act utilitarianism and rule utilitarianism. Mill does not distinguish between these two forms in his writing, and different passages suggest different interpretations (the difference probably wasn’t clear in his own mind).

*Act utilitarianism* is thought to be the “pure” utilitarian position, where each act is considered on its own merits. For any particular act, if performing it will maximize the good, then it should be performed; otherwise not. Act utilitarianism may lead to certain theoretical problems (such as urging us not to keep private promises), for which reason some ethicists have promoted a modified version of utilitarianism called “rule utilitarianism.”

With *rule utilitarianism*, the item of moral evaluation isn’t the individual act, but rather the rule it follows: if following a certain rule (instead of some other rule) maximizes the good, then that rule should be followed, even if it would turn out, with some instances, that happiness could be maximized by breaking the rule. This means, for instance, that certain applications of the rule might fail to maximize the good, but because that *kind* of act normally does maximize the good, then it is always right to so act. This form of utilitarianism has the advantage of being easier for human beings to follow: we have to evaluate

only rules, rather than individual acts. It also has the advantage of avoiding certain problems of act utilitarianism, such as committing unjust (yet happiness-maximizing) acts. It has the intuitive disadvantage, however, of occasionally requiring us to perform acts even when doing so will fail to maximize the good.

## SOME MISUNDERSTANDINGS THAT MILL ADDRESSES

### Utilitarianism is a Swine's Morality

Critics of utilitarianism have claimed that it is simply a kind of hedonism, equating what is morally good with the sensation of pleasure, and that this is no different than what swine pursue, namely, their own pleasure of rolling in the mud and swilling at the trough. The gist of this criticism is that Mill has misunderstood human nature: he believes that human beings desire only to “eat, drink, fornicate, and snore” (to quote one critic), when in fact they desire far more than this.

In his response to this criticism, Mill agrees that humans desire far more than these bodily pleasures, but chides the critics for assuming that this is the *only* kind of pleasure. Mill finds that there are *intellectual* as well as bodily pleasures, and that the former are even more desirable than the latter. What Mill seems to be claiming here is not that these are distinct *feelings*, but rather that we have *different faculties* that are gratified, each resulting in a feeling of pleasure. In other words, we do not obtain pleasure in all the same ways as a pig; rather, we have certain faculties that cannot be gratified by swinish behavior.

Mill gives two arguments for the desirability of the higher pleasures. First, intellectual pleasures afford greater “permanency, safety, uncostliness.” Second, some kinds of pleasure are simply more desirable and more valuable than others, and these pleasures can be so ranked. How do we know this? We appeal to the “*Millian judge*,” someone who has known both kinds of pleasure: the objective means for ranking these pleasures according to their desirability is to ask such a judge which is preferable. And the answer, from those who have known both bodily and intellectual pleasures, is that the latter is far more desirable.

### Utilitarianism is “Too Low a Calling” for Humans

Here the complaint is that utilitarianism doesn't ask enough of us. All it asks is that we do what makes us happy — but it seems as though we will do that anyway. Morality ought to be more challenging, it ought to encourage us to lead “better” lives, and utilitarianism does not seem to do this.

Mill replies that “in the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility.” What he apparently means by this is that the **impartiality** built into the Golden Rule (namely, that we should treat others as we would want them to treat us) is also included in the Greatest Happiness Principle. And indeed it is. The GHP does not instruct me to maximize my own happiness, but rather to maximize the overall happiness of all “sentient creation” (that is, all beings capable of feeling pleasure or pain). That means that I might often be required to sacrifice my own happiness (perhaps even my own life) so as to maximize the total happiness. What is more, utilitarianism offers suggestions for promoting *compliance* with this impartiality requirement (in the sanctions mentioned above).

# READING

---

## *UTILITARIANISM* (SELECTION)

John Stuart Mill

---

*John Stuart Mill (1806-73) was born in London as the first son of the Scottish philosopher James Mill. The elder Mill personally undertook the education of John Stuart and, as a consequence of certain natural gifts, and perhaps also of the pedagogy employed, the young student was reading Latin by the age of three and Greek by the age of eight — and in*

*general was well versed in the arts and sciences by the time he was in his teens.*

*Mill went to work for the British East India Company at the age of seventeen, and stayed in that employment for thirty-five years. In 1852 he married Harriet Taylor, a recently widowed woman with whom he had shared an in-*

tense, intellectual relationship for the previous twenty-one years. He was elected to Parliament for a term in 1865.

Both during and after his employment with the East India Company, Mill made important contributions to philosophy and social reform, writing such classics as *A System of Logic* (1843), *Principles of Political Economy* (1848), *On Liberty* (1859), *Utilitarianism* (1863), and *The Subjection of Women* (1869). The following is a selection from chapter two of *Utilitarianism*.

---

The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure. To give a clear view of the moral standard set up by the theory, much more requires to be said; in particular, what things it includes in the ideas of pain and pleasure; and to what extent this is left an open question. But these supplementary explanations do not affect the theory of life on which this theory of morality is grounded — namely, that pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain.

Now, such a theory of life excites in many minds, and among them in some of the most estimable in feeling and purpose, inveterate dislike. To suppose that life has (as they express it) no higher end than pleasure — no better and nobler object of desire and pursuit they designate as utterly mean and groveling; as a doctrine worthy only of swine, to whom the followers of Epicurus were, at a very early period, contemptuously likened; and modern holders of the doctrine are occasionally made the subject of equally polite comparisons by its German, French, and English assailants.

When thus attacked, the Epicureans have always answered, that it is not they, but their accusers, who represent human nature in a degrading light; since the accusation supposes human beings to be capable of no pleasures except those of which swine are capable. If this supposition were true, the charge could not be gainsaid, but would then be no longer an imputation; for if the sources of pleasure were precisely the same to human beings and to swine, the rule of life which is good enough for the one would be good enough for the other. The comparison of the Epicurean life to that of beasts is felt as degrading, precisely because a beast's pleasures do not satisfy a human being's conceptions

of happiness. Human beings have faculties more elevated than the animal appetites, and when once made conscious of them, do not regard anything as happiness which does not include their gratification. I do not, indeed, consider the Epicureans to have been by any means faultless in drawing out their scheme of consequences from the utilitarian principle. To do this in any sufficient manner, many Stoic, as well as Christian elements require to be included. But there is no known Epicurean theory of life which does not assign to the pleasures of the intellect, of the feelings and imagination, and of the moral sentiments, a much higher value as pleasures than to those of mere sensation. It must be admitted, however, that utilitarian writers in general have placed the superiority of mental over bodily pleasures chiefly in the greater permanency, safety, uncostliness, etc., of the former — that is, in their circumstantial advantages rather than in their intrinsic nature. And on all these points utilitarians have fully proved their case; but they might have taken the other, and, as it may be called, higher ground, with entire consistency. It is quite compatible with the principle of utility to recognise the fact, that some kinds of pleasure are more desirable and more valuable than others. It would be absurd that while, in estimating all other things, quality is considered as well as quantity, the estimation of pleasures should be supposed to depend on quantity alone.

If I am asked, what I mean by difference of quality in pleasures, or what makes one pleasure more valuable than another, merely as a pleasure, except its being greater in amount, there is but one possible answer. Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure. If one of the two is, by those who are competently acquainted with both, placed so far above the other that they prefer it, even though knowing it to be attended with a greater amount of discontent, and would not resign it for any quantity of the other pleasure which their nature is capable of, we are justified in ascribing to the preferred enjoyment a superiority in quality, so far outweighing quantity as to render it, in comparison, of small account.

Now it is an unquestionable fact that those who are equally acquainted with, and equally capable of appreciating and enjoying, both, do give a most marked preference to the manner of existence which employs their higher faculties. Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an igno-

ramus, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs. They would not resign what they possess more than he for the most complete satisfaction of all the desires which they have in common with him. If they ever fancy they would, it is only in cases of unhappiness so extreme, that to escape from it they would exchange their lot for almost any other, however undesirable in their own eyes. A being of higher faculties requires more to make him happy, is capable probably of more acute suffering, and certainly accessible to it at more points, than one of an inferior type; but in spite of these liabilities, he can never really wish to sink into what he feels to be a lower grade of existence. We may give what explanation we please of this unwillingness; we may attribute it to pride, a name which is given indiscriminately to some of the most and to some of the least estimable feelings of which mankind are capable; we may refer it to the love of liberty and personal independence, an appeal to which was with the Stoics one of the most effective means for the inculcation of it; to the love of power, or to the love of excitement, both of which do really enter into and contribute to it; but its most appropriate appellation is a sense of dignity, which all human beings possess in one form or other, and in some, though by no means in exact, proportion to their higher faculties, and which is so essential a part of the happiness of those in whom it is strong, that nothing which conflicts with it could be, otherwise than momentarily, an object of desire to them.

Whoever supposes that this preference takes place at a sacrifice of happiness — that the superior being, in anything like equal circumstances, is not happier than the inferior — confounds the two very different ideas, of happiness, and contentment. It is indisputable that the being whose capacities of enjoyment are low, has the greatest chance of having them fully satisfied; and a highly endowed being will always feel that any happiness which he can look for, as the world is constituted, is imperfect. But he can learn to bear its imperfections, if they are at all bearable; and they will not make him envy the being who is indeed unconscious of the imperfections, but only because he feels not at all the good which those imperfections qualify. It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are a different opinion, it is because they only know their own side of the question. The other party to the comparison knows both sides. [...]

I have dwelt on this point, as being a necessary part of a perfectly just conception of Utility or Happiness, considered as the directive rule of human conduct. But it is by no means an indispensable condition to the acceptance of the utilitarian standard; for that standard is not the agent's own greatest happiness, but the greatest amount of happiness altogether; and if it may possibly be doubted whether a noble character is always the happier for its nobleness, there can be no doubt that it makes other people happier, and that the world in general is immensely a gainer by it. Utilitarianism, therefore, could only attain its end by the general cultivation of nobleness of character, even if each individual were only benefited by the nobleness of others, and his own, so far as happiness is concerned, were a sheer deduction from the benefit. But the bare enunciation of such an absurdity as this last, renders refutation superfluous.

According to the Greatest Happiness Principle, as above explained, the ultimate end, with reference to and for the sake of which all other things are desirable (whether we are considering our own good or that of other people), is an existence exempt as far as possible from pain, and as rich as possible in enjoyments, both in point of quantity and quality; the test of quality, and the rule for measuring it against quantity, being the preference felt by those who in their opportunities of experience, to which must be added their habits of self-consciousness and self-observation, are best furnished with the means of comparison. This, being, according to the utilitarian opinion, the end of human action, is necessarily also the standard of morality; which may accordingly be defined, the rules and precepts for human conduct, by the observance of which an existence such as has been described might be, to the greatest extent possible, secured to all mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation. [...]

I must again repeat, what the assailants of utilitarianism seldom have the justice to acknowledge, that the happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator. In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility. To do as you would be done by, and to love your neighbor as yourself, constitute the ideal perfection of utilitarian morality. [...]

# DEONTOLOGICAL ETHICS

---

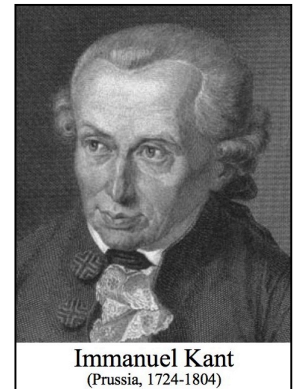
---

“FROM THE CROOKED TIMBER OF HUMANITY,  
NOTHING WHOLLY STRAIGHT CAN BE MADE.”

– Immanuel Kant (1724-1804)

## DUTY AND THE GOOD WILL

Immanuel Kant (1724-1804) wrote prolifically on a wide range of subjects, most famously on epistemology and the limits of human reason in his *Critique of Pure Reason* (1781). But he is also well known for his moral philosophy, and here he wants to explain the feeling that many of us have that certain actions are required or prohibited of us *absolutely*, unconditionally, without exception. For instance, many people feel that they are absolutely prohibited from torturing or killing innocent human beings *no matter what*, even if the whole world depended upon it. The source of this obligation cannot be our desire for some consequence of our action or omission, for we find the action or omission to be right or wrong *in itself*, independent of the consequences. I simply see that it is wrong to take innocent life, and I avoid it not because I fear going to jail, nor because I fear public criticism, nor because I am merely disinclined to kill the innocent, but because it is morally wrong.



Now, what could be the source of such an unyielding sense of obligation? One might think it is simply a strong, emotion-laden inclination based on previous conditioning and perhaps our biological nature. But Kant suggests it has instead to do with the nature of reason itself, that reason is the source of this obligation and feeling of duty.

Kant discovers in our reason a moral principle called the Categorical Imperative, which he uses to discover more specific, lower-level moral laws or *duties* (also called “categorical imperatives” or “moral imperatives” or “imperatives of duty”), and it becomes our self-imposed duty to follow these moral laws. Kant also discovered a logical difference between two different kinds of duties — what he calls perfect and imperfect duties — and this difference is also of moral and social importance, for perfect duties appear to be the necessary conditions for human existence within any society, while imperfect duties are the necessary conditions for human existence within any society that is worth having (or “is desirable”).

## INTENTIONS, NOT CONSEQUENCES

Kant’s ethics emphasizes the motives and intentions of a person’s actions rather than the consequences. The will, for Kant, is that capacity found in human beings for acting from a principle. As Kant writes:

Everything in nature works according to law. Only a rational being has the capacity of acting according to the conception of laws, that is, according to principles. This capacity is will. [Ak. 4: 412]

Insofar as the human being guides her actions according to some maxim or principle, rather than according to some whim or inclination — in other words, when an action is motivated by some principle — then the human being is a *person*, possessing dignity and worth, and is more than a mere biological creature. Only human beings are capable of moral good because only they have **reason** (the ability to conceive of alternative possibilities) and **freedom** (the ability to choose and act on these possibilities); but unless they use this reason and freedom, these human beings are not persons, and are not much better off than cattle. Kant characterizes this point quite nicely in a passage from his *Metaphysics of Morals* (1797):

In the system of nature, a human being is a being of slight importance and shares with the rest of animals, as offspring of the earth, an ordinary value. Although a human being has, in his understanding, something more than they and can set himself ends, even this gives him only an *extrinsic* value for his usefulness; that is to say,

it gives one man a higher value than another, that is, a *price* as of a commodity in exchange with these animals as things....

But a human being regarded as a *person*, that is, as the subject of a morally practical reason, is exalted above any price; for as a person he is not to be valued merely as a means to the ends of others or even to his own ends, but as an end in itself, that is, he possesses a dignity (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. [Ak. 6: 434-5, Gregor translation (Latin phrases omitted)]

What confers worth upon us is the exercising of our wills, our acting according to some maxim or principle. This elevates us above mere cattle, which chew their cud as a matter of inclination. Furthermore, when our action is based on the *right* maxim or principle, then the will is *morally good*. To act from inclination, however, is not to act with one's will at all — here the will is simply idling. This concept of the will is central to Kant's theory.

## WHAT IS THE SOURCE OF VALUE?

If something is valued as an end, for its own sake, then it has *final value*. It is the ultimate reason for all human action. Aristotle, Kant, and Mill — and pretty much everyone else — believe that happiness is what has final value. What has *instrumental value*, on the other hand, is anything valued because it brings about happiness.

Final and instrumental value both concern the *way* that we value something. Having decided this, however, we might still ask about the *source* of value. Whatever is the source of value we call intrinsically valuable, while everything else of value will be valuable only so far as it is related to the source of value in the right way (and so is valuable only extrinsically). For Kant, the good will is the source of value, and happiness has value only if it is associated with the good will. Similarly, anything that a good will desires will be thought to have value simply because the good will desires it.

Mill and other utilitarians view happiness as the source of value, as well as having final value. Kant realized that all human beings desire happiness, and that we desire it for itself; but Kant also believed that happiness, apart from a good will, was without moral value. If happiness were the source of value, then it wouldn't matter how we obtained it, but as the common saying goes, "the end does not justify the means," or at least not always, and some instances of happiness strike nearly everyone as not merely void of value, but positively bad. Imagine, for instance, some brutal, wealthy fellow who amassed his fortunes by exploiting illegal aliens in slavery-like conditions in his sweatshop, and suppose that this wealth brought him great happiness. Kant claims that such happiness will always strike an impartial observer as being without value.

Kant begins section one of the *Foundations* with his memorable claim that "nothing in the world ... can possibly be conceived which could be called good without qualification except a good will." The good will has **intrinsic value** ("has absolute value," "is good in itself"), and all other value is derived from this source, and so is **extrinsic**. *Happiness* can't have intrinsic value, for while we are all naturally inclined to seek happiness, Kant notes that we do not consider as good those who enjoy happiness at the expense of others. We view such people "with disapproval."<sup>3</sup> The happiness felt by a successful sadistic murderer has no value (as judged by an "impartial observer") because it is not joined with a good will. Kant's main concern is not with "what makes me happy" but rather with "what makes me worthy of happiness" — and having a good will seems necessary for a person to be considered worthy of happiness. For John Stuart Mill, *happiness* is the source of value, the end-in-itself. For Kant, *persons* are the source of value. A corollary of this: for Mill, we might sacrifice people so as to maximize happiness; for Kant, we would instead sacrifice happiness for the sake of persons.

## MOTIVATIONS FOR ACTING: INCLINATION AND DUTY

There are two possible motivations for any of our actions: *inclination* (which includes our sensuous or animal desires) and *duty* (the voice of reason). As for inclination, we might be inclined by the desire for some *consequence* of the action (these are so-called non-immediate inclinations), or we might be inclined by a desire for the action itself (so-called immediate inclinations). Were an action motivated by duty, however, the agent would perform the action even if the consequences were not desired and the action itself was repugnant. The "call of duty" is the feeling that you must do something *not* because of some desired or feared consequence, nor because you want to do it, but because it is right, because it is the thing to do and so you

<sup>3</sup> It is hard not to think here of that often-quoted passage from the Christian scriptures, I Corinthians 13, where the author writes: "If I have all the eloquence of men or of angels, but speak without love, I am simply a gong booming or a cymbal clashing... (etc)." Here love (Greek: *agape*) is seen in a way analogous to Kant's good will.



must do it. Even if the world were to be destroyed tomorrow, you would still feel obliged to honor this duty. The most compelling examples are of duties *not* to act in certain ways, such as the duty not to torture children — recall Alyosha Karamazov, who would not torture a child even if it were to bring about universal happiness for the rest of humanity.

This notion of duty is built into the word ‘deontological’, the stem of which comes from the Greek ‘deon’ [= that which is binding or needful] or ‘dein’ [= to bind]. Similarly, ‘obligation’ comes from the Latin ‘ob’ [= in the way of; towards or against] and ‘ligare’ [= to bind]. We speak of being “duty bound,” for instance. Duty and obligation are closely related concepts, and are central in Kant’s deontological ethics.

### THE GOOD WILL IS A WILL ACTING FROM DUTY

To have a good will means that one is acting from duty. Reason, which is the same for everyone, determines those actions we have a duty to perform; if we then choose those actions because we see that they are our duty, then we are acting from a good will. Our will is considered good simply because of its motivations; the consequences of our actions are irrelevant to the will’s value. To say that the good will has intrinsic value is to say that it is good in itself, independent of all else and any possible consequences its actions might have – its value does not depend on it being a means to another good, such as happiness.

The utilitarian judges the rightness of actions solely on the basis of their consequences. For Kant, an action is right so long as it *conforms* with duty. Kant then adds a new distinction, claiming that the action might be morally right, but it still lacks moral value (and the will choosing will not be considered good) unless the action is also *motivated* by duty. For Kant, none of the possible consequences of an action can serve as a motive (e.g., where the self *desires* some likely consequence and therefore performs the action that is hoped to produce it). The agent’s motive must be *duty*. In contrast, the utilitarian considers only the consequences in assessing an action.

### ACTIONS IN CONFORMITY WITH DUTY AND ACTIONS MOTIVATED BY DUTY

Kant offers five illustrations of this three-fold distinction of motive (non-immediate and immediate inclination, and duty): (1) the honest grocer, (2) preserving one’s life, (3) helping others where one can, (4) assuring one’s own happiness, (5) loving neighbors and enemies.

Let’s consider the honest grocer example. Imagine some grocer who treats all of his customers fairly and does not take advantage of young children or the feeble-minded, or foreigners unfamiliar with the local currency. Clearly such honesty is in conformity with the grocer’s duty (we’ll investigate *why* it is later). But the motivation behind such action might not be duty itself, but instead some inclination. For instance, some actions conform to duty but are done from a “non-immediate inclination.” Suppose the grocer’s honesty is motivated by a desire for its good consequences, namely, continued patronage of his store – or conversely, a desire to avoid the eventual consequences of dishonesty. To do something out of a desire for the consequences of the act is to act from a non-immediate inclination. Here the grocer is being honest not because he sees that it is his duty, but because he desires the rewards of honesty or fears the penalties of dishonesty.

Other actions conform to duty, but are done from some “*immediate*” *inclination*. Suppose that the grocer’s honesty is motivated by a natural affection for his customers. Suppose he is of that agreeable nature that rejoices in helping others, and that cheating a customer would be the farthest thing from his mind. Here the honest behavior is itself desired. Whenever an action (or omission) is motivated by a desire or abhorrence of the action itself (as opposed to any consequences of the action), then the action is motivated by an “immediate inclination.”

Finally, some actions both conform to duty and are motivated by duty. If the grocer’s honesty is motivated by the realization that honest behavior is his duty, then his action *now* has moral value. The good will is the source of value, and it is defined as a will that acts from duty; so only such actions have any value.<sup>4</sup>

---

<sup>4</sup> Note that some actions conform with duty *yet are trivial* or morally irrelevant, for example, writing your name in your books or pulling your left sock on before your right sock. These are not done *from respect for the law*, but only in conformity with it, and so lack moral worth.

## DUTY AND IMPERATIVES

Kant defines a good will as one that is motivated by duty; but how do we decide what our duty is? Our duty will be those imperatives that bind us categorically or absolutely, and these can be found to bind us categorically because of their logical form (or, alternatively, because of the nature of human beings as the ultimate source of value). So we will first consider what categorical imperatives are, and then examine their logical form (as displayed in the first formulation of the Categorical Imperative) and how they are related to humanity as the source of value (in the second formulation of the Categorical Imperative).

### HYPOTHETICAL AND CATEGORICAL IMPERATIVES

An imperative is simply a command, and Kant distinguishes between two general kinds: hypothetical (of which there are imperatives of skill and imperatives of prudence) and categorical. Only the latter are absolutely binding. Hypothetical imperatives all have the logical form: “If you want E, then do A!”, where E is some particular end or goal, and A is some action: with imperatives of skill, the thing wanted might be any goal at all, while an imperative of prudence always posits happiness as the goal – a goal apparently shared by all humans. Categorical imperatives, on the other hand, simply have the form: “Do A!”

Kant wants to explain and account for our “sense of duty,” which is similar to the voice of conscience commanding us to do (or *not* do) something. What is the nature of this command? Is it hypothetical? Seemingly not. It binds us absolutely, and so is categorical. With hypothetical imperatives an action is required to attain some end, while with categorical imperatives an action is simply required in and of itself.

Most of the imperatives that we hear in life are hypothetical, although they are often uttered *as if* they were categorical (that is to say, their true logical form is hypothetical, although their surface or apparent form is categorical). For instance, a mother might command her child: “Eat your peas!” It looks as though the mother is issuing a categorical command, and yet no one would mistake this as a moral utterance; it would be a strange world in which children had a moral duty to eat their peas. Rather, the true form of her command is hypothetical, and she has simply left off the antecedent part: she’s really saying “If you want to have any dessert, then eat your peas!” or “If you want to leave the table, then eat your peas!” or “If you want to be healthy, then eat your peas!” and so on.

Hypothetical imperatives are not absolutely binding because they always assume some desired end, and so the command can be avoided simply by rejecting that end. The child can always reply to her mother: “I don’t need to eat those peas because I don’t care for any dessert” (or “... because I don’t care if you beat me” or “... because I don’t care about my physical health”). The example of the peas is an imperative of skill: these always refer to an end that you might *possibly* want, and thus are commanded to act in a way that will not foreclose your ability to obtain that end. But because the end is always contingent, the command is contingent as well.

One might think that imperatives of prudence, on the other hand, are inescapable, since they posit one’s own happiness as the end of some action (“If you want to be happy, then do A!”). Kant agreed with Aristotle and others that happiness is an end towards which all human beings strive, and so it might stand to reason that such imperatives would indeed be absolutely binding on us, thus serving as a basis for our moral claims. Yet closer inspection reveals that such claims still lack an ultimate binding force. I could with all sincerity encourage you to study more philosophy with the claim: “If you want to be happy, then study philosophy!” – for although philosophy is often difficult, and occasionally frustrating, it will make possible for you degrees and kinds of happiness that you aren’t likely to find elsewhere. Nevertheless, many of you will manage not to believe me, and so will avoid my command by pointing out that, while you *do* want to be happy, you don’t think studying philosophy will get you there. Imperatives of prudence are therefore still hypothetical, for although everyone desires happiness as an end, what actually brings about happiness for each person may be different.

Categorical imperatives are imperatives of morality, and have the form: “Do A.” Here an action is required in and of itself, regardless of any possible ends. It is this species of command that Kant is trying to identify and explain with his moral theory.

## MORALITY AND RELIGION

Many people base – or at least believe that they base – their moral views upon their religious views. A typical notion is that we are to act morally because of what might happen after we die: if we act immorally (i.e., if we sin) then we will roast in hell eternally; if we act morally (i.e., do God’s bidding) we will gain entrance to heaven and its eternal rewards. But to base morality on religion like this is to make the commands of morality all *hypothetical*, having the general form: “If you want to get to heaven (or avoid hell), then do A!” This turns morality into little more than a kind of prudential reasoning, and because Kant believes that all true moral claims are categorical, he rejects this sort of religious foundation. Nor does his rejection seem entirely implausible. After all, how many of us refrain from murdering others, or torturing children, or littering, simply because we fear that not doing so might jeopardize our afterlife? Don’t we believe that it’s wrong to torture children *regardless of the consequences*?

## THE BASIS OF MORAL OBLIGATION

To what extent do these hypothetical and categorical imperatives *bind* us, that is, oblige us? Imperatives of skill have little binding power since we need merely say that we don’t desire the end. Imperatives of prudence would appear to be more binding, since everyone desires happiness; but the path to happiness may be different for people, and so the imperative may simply be wrong. Imperatives of morality, on the other hand, bind us completely – and this accords well with how we feel about duty.

At the same time, why hypothetical imperatives bind us is transparently clear: insofar as we desire the end, then we are bound to perform the necessary means. Thus imperatives of skill are tautologically true: you want whatever means are necessary to some desired end. Imperatives of prudence are also tautological, as long as you really know what means *are* necessary for happiness – but generally these are more like *counsels* than *commands*: “doing X will *tend* towards happiness.” Both of these imperatives are based on some *desire*: if you desire something, then you must do whatever is necessary to obtain it, and the obligatory force of the imperative depends on that action being a necessary means to the satisfaction of the desire.

Imperatives of morality are *not* based on the desire for some end, and the ultimate source of their hold upon us is that they arise from the reason within each one of us. Reason discovers moral laws within itself just as it discovers physical laws within the world, and so it is to these that we must now turn.

## RESPECT FOR THE LAW

To act from duty means to act out of respect for the moral law. Here the “moral law” is seen as analogous to a “physical law”: it is a rule of action that all humans are to follow. These rules are called categorical imperatives, or imperatives of morality or duty, and they are generated by what Kant calls *the* Categorical Imperative, which reads (in its first formulation):

**“Act only on that maxim that I can consistently will to become a universal law”**

This is a rule for telling us how to make rules of action (a rule for rule making). At its heart is the prohibition against making a moral exception of oneself (there is to be “no double-standard” – one for me and another for everyone else). To understand this criterion we need to understand its parts, namely, what maxims and universal laws are, and what is involved in “consistently willing” something.

A maxim is a subjective action-guide, a principle for guiding my actions in particular situations. For example:

- “I shall treat other humans as mere means to my own happiness.”
- “I shall cheat on my exams if it will help my GPA.”
- “I shall help others when I can without serious risk to myself.”
- “I shall steal any library books that I really need.”

A universal law, on the other hand, is an *objective* action-guide, that is, a principle for guiding *everyone*’s actions. For example:

- “Everyone shall treat other humans as mere means to their own happiness.”
- “Everyone shall cheat on their exams if it helps their GPA.”
- “Everyone shall help others if it involves no serious risk to themselves.”
- “Everyone shall steal any library books that they really need.”

It turns out that there are two applicable senses of “willing consistently” that a maxim be made into a universal law, and these two senses distinguish what Kant calls perfect and imperfect duties. With perfect (or “strict”) duties, the opposite maxim is *logically impossible* (i.e., self-contradictory) when universalized (see the second example, below). With imperfect (or “meritorious”) duties, the opposite maxim *can* be universalized, but I cannot *want* to universalize it (see the fourth example, below). In other words, some maxims – when universalized – result in a law that becomes self-defeating (these are *logically* inconsistent, and the opposite maxim is for us a perfect duty). Other maxims *can* be universalized, but result in a situation that we do not really want (thus result in a contradiction in our will, showing that the opposite maxim is for us an imperfect duty).

If it was from duty that I did not steal the book, then what motivated my honesty in the matter was my recognition that the *maxim* governing such an action could not be consistently willed to be a *law* (a rule governing everyone’s actions, including my own).

## KANT’S EXAMPLES OF PERFECT AND IMPERFECT DUTIES

Because there are duties to oneself as well as duties to others, and because there are perfect as well as imperfect duties, Kant offers us four sample duties: a perfect duty to the self, a perfect duty to others, an imperfect duty to the self, and an imperfect duty to others (Ak. 4: 421-23). There are, of course, many other duties that fall into these four categories.

### (1) Perfect duty to self: “From self-love end your life if it minimizes pain”

Kant believes that I have a perfect duty not to commit suicide, because I cannot universalize the opposite maxim. This “law” of ending one’s life out of self-love is inconsistent with itself; it cannot be universalized because the same principle of behavior (self-love) cannot, in a rational system, lead to diametrically opposed behaviors (viz., the furtherance of one’s life and the destruction of one’s life). This makes sense primarily when the moral world is considered as analogous with the physical world, where the same natural law cannot issue in opposite behaviors. Think of moral laws as natural human instincts: Here we have an instinct for survival (“self-love”), and it would be contradictory that this same instinct also desire its opposite. This does *not*, however, prohibit all suicide (such as killing oneself to help another: the morality of this would require further inquiry); it only prohibits suicide motivated by self-love.

### (2) Perfect duty to others: “Make false promises when convenient”

I have a perfect duty not to make false promises, because I *cannot* universalize the opposite maxim. Universalization is not possible because it is logically inconsistent with the very institution of promise-making.

Kant is *not* saying here that I must not give false promises because *eventually* the institution of promise-keeping would be undermined, and that I do not want such an institution (this would result in a merely prudential, or hypothetical imperative [Why not an imperfect duty?]). Rather, the universalization of the maxim results *immediately* in a logical contradiction.

### (3) Imperfect duty to self: “Let your talents rust if you are content”

I have an imperfect duty to develop my talents, because I *cannot will* to universalize the opposite maxim. This law is inconsistent with my will. We can *universalize* this maxim (making idleness even an *instinct* in us), but we cannot *will* that it be so universalized, for our ends are often changing, and we always desire the means necessary to attain the end. In the future, my ends may be such that I will have required the cultivation of those very talents that I am now neglecting. So there is a contradiction in my will: I will that I *do not* cultivate my talents and I will that I *do* cultivate my talents (in order to attain possible future ends).

### (4) Imperfect duty to others: “Let all people fend for themselves”

I have an imperfect duty to help others in need, because I *cannot will* to universalize the opposite maxim. This cannot be universalized because no matter how well off I may be now, there is always the possibility that I will some day be in need of help. Because I will then desire the aid of others, I cannot *also* desire that no one ever give aid. This would involve a contradiction in my will.

### THREE FORMULATIONS – ONE IMPERATIVE

This one Categorical Imperative is meant as a *principle* or *formula* from which is to be derived all the commands of morality (the laws that practical reason gives for guiding our actions). So it is really a second-order rule of action, from which we derive our first-order rules (e.g., of promise keeping, or of benevolence). Kant also gives us three separate formulations of this imperative – the Universal Law, the End-in-Itself, and the Autonomy formulations – noting that these are three ways of thinking about the same thing, and not different moral principles.

The Universal Law formulation (“Act only on that maxim which you can consistently will to be a universal law” [Ak. 4: 421]) was discussed above with Kant’s four examples and the discussion of perfect and imperfect duties.

The End-in-Itself formulation (“Act so as to treat humanity, whether in your own person or in the person of another, always as an end, never merely as a means” [Ak. 4: 429]) focuses on the nature of human beings, insofar as they act according to maxims (“have wills”) and so are persons. According to this formulation, I am not to use another in any way with which the other cannot in principle agree, since doing so would be to use that person *merely* as a means, as a mere tool or instrument of my own plans and desires. This forbids the use of deception or coercion, since either of these involve the other person in a scheme of action to which they would not consent if they knew all the details (were not deceived) or if they were not forced. For instance, to make a lying promise to another so as to procure a loan is to use the person lied to as a mere means. One might as well hit him over the head and steal the money outright – the difference here between force and fraud is morally negligible.

Also, we act for various ends, most of which are relative, and so differ from person to person. But is there an end of absolute value? If so, it is an end common to *all* humans, and so can be the basis of a common principle of action. Well, one end we all share is happiness – even Kant believes this (since happiness is a final good). But happiness doesn’t exist in itself; it can occur only in humans, so we are to promote one another’s happiness, which is to promote one another’s ends, since having one’s ends promoted results in happiness. *Perfect duties* require that we not treat a person as a mere means. *Imperfect duties* require that we promote the interests of others (we view another as an end in itself, and therefore wish also to promote its interests).

The Autonomy formulation (“Act so that the will may regard itself as in its maxims laying down universal laws” [Ak. 4: 431]) instructs us to act as autonomous agents legislating for all agents in the kingdom of ends. Everyone is legislating for themselves, and at the same time for everyone else, in that we are all using the same basic formula for deciding which of our maxims are moral, and which not.

## READING

---

### *FOUNDATIONS OF THE METAPHYSICS OF MORALS* (SELECTION)

Immanuel Kant

---

*Immanuel Kant (1724-1804) was born, raised, lived, and died in Königsberg (East Prussia, later named Kaliningrad as part of the former USSR). Kant was the first modern philosopher to teach philosophy in a university — and after Plato, Aristotle, and maybe Descartes, he has done the most to alter the way that philosophers pursue their discipline. He was “awakened from his dogmatic slumbers” rather late in life by reading David Hume, and went on to write one of the greatest (and most difficult) books in the history of philosophy: The Critique of Pure Reason (1781; 2nd ed.: 1787). While at first widely misunderstood, this book went*

*on to change the way that we think of ourselves and the physical universe.*

*As a young professor, Kant was quite the socialite, always in high demand at parties and other gatherings. In his later years, however, his life became more regular, and legend claims the neighbor women would set their clocks by his afternoon walks, which he began promptly at 3:30 (the path that he took came to be called “The Philosopher’s Walk”). Only once did Kant fail in this routine: Having recently received Rousseau’s new book *Émile*, he was unable to tear himself away from it.*

*The following reading is from sections one and two of Kant's Foundations of the Metaphysics of Morals (published in 1785), in which Kant formulated for the first time the general outlines of his new moral theory.*

---

**[THE GOOD WILL]**

Nothing in the world — indeed nothing even beyond the world — can possibly be conceived which could be called good without qualification except a *good will*. Intelligence, wit, judgment, and the other talents of the mind, however they may be named, or courage, resoluteness, and perseverance as qualities of temperament, are doubtless in many respects good and desirable. But they can become extremely bad and harmful if the will, which is to make use of these gifts of nature and which in its special constitution is called character, is not good. It is the same with the gifts of fortune. Power, riches, honor, even health, general well-being, and the contentment with one's condition which is called happiness, make for pride and even arrogance if there is not a good will to correct their influence on the mind and on its principles of action so as to make it universally conformable to its end. It need hardly be mentioned that the sight of a being adorned with no feature of a pure and good will, yet enjoying uninterrupted prosperity, can never give pleasure to a rational impartial observer. Thus the good will seems to constitute the indispensable condition even of worthiness to be happy.

[...]

The good will is not good because of what it effects or accomplishes or because of its adequacy to achieve some proposed end; it is good only because of its willing, i.e., it is good of itself. And, regarded for itself, it is to be esteemed incomparably higher than anything which could be brought about by it in favor of any inclination or even of the sum total of all inclinations. Even if it should happen that, by a particularly unfortunate fate or by the niggardly provision of a stepmotherly nature, this will should be wholly lacking in power to accomplish its purpose, and if even the greatest effort should not avail it to achieve anything of its end, and if there remained only the good will (not as a mere wish but as the summoning of all the means in our power), it would sparkle like a jewel in its own right, as something that had its full worth in itself. Usefulness or fruitlessness can neither diminish nor augment this worth. [...]

**[THE GOOD WILL AND DUTY]**

We have, then, to develop the concept of a will which is to be esteemed as good of itself without regard to anything else. It dwells already in the natural sound understanding and does not need so much to be taught as only to be brought to light. In the estimation of the total worth of our actions it always takes first place and is the condition of everything else. In order to show this, we shall take the concept of duty. It contains that of a good will, though with certain subjective restrictions and hindrances; but these are far from concealing it and making it unrecognizable, for they rather bring it out by contrast and make it shine forth all the brighter.

**[ACTIONS MOTIVATED BY DUTY]**

I here omit all actions which are recognized as opposed to duty, even though they may be useful in one respect or another, for with these the question does not arise at all as to whether they may be carried out *from* duty, since they conflict with it. I also pass over the actions which are really in accordance with duty and to which one has no direct inclination, rather executing them because impelled to do so by another inclination. For it is easily decided whether an action in accord with duty is performed from duty or for some selfish purpose. It is far more difficult to note this difference when the action is in accordance with duty and, in addition, the subject has a direct inclination to do it. For example, it is in fact in accordance with duty that a dealer should not overcharge an inexperienced customer, and wherever there is much business the prudent merchant does not do so, having a fixed price for everyone, so that a child may buy of him as cheaply as any other. Thus the customer is honestly served. But this is far from sufficient to justify the belief that the merchant has behaved in this way from duty and principles of honesty. His own advantage required this behavior; but it cannot be assumed that over and above that he had a direct inclination to the purchaser and that, out of love, as it were, he gave none an advantage in price over another. Therefore the action was done neither from duty nor from direct inclination but only for a selfish purpose. [...]

**[THE WILL AND THE LAW]**

Everything in nature works according to laws. Only a rational being has the capacity of acting according to the conception of laws, i.e., according to principles. This capacity is will. Since reason is required for the derivation of

actions from laws, will is nothing else than practical reason. If reason infallibly determines the will, the actions which such a being recognizes as objectively necessary are also subjectively necessary. That is, the will is a faculty of choosing only that which reason, independently of inclination, recognizes as practically necessary, i.e., as good. [...]

#### [CLASSIFICATION OF IMPERATIVES]

The conception of an objective principle, so far as it constrains a will, is a command (of reason), and the formula of this command is called an *imperative*.

All imperatives are expressed by an “ought” and thereby indicate the relation of an objective law of reason to a will which is not in its subjective constitution necessarily determined by this law. This relation is that of constraint. Imperatives say that it would be good to do or to refrain from doing something, but they say it to a will which does not always do something simply because it is presented as a good thing to do. Practical good is what determines the will by means of the conception of reason and hence not by subjective causes but, rather, objectively, i.e., on grounds which are valid for every rational being as such. It is distinguished from the pleasant as that which has an influence on the will only by means of a sensation from merely subjective causes, which hold only for the senses of this or that person and not as a principle of reason which holds for everyone.

[...]

All imperatives command either hypothetically or categorically. The former present the practical necessity of a possible action as a means to achieving something else which one desires (or which one may possibly desire). The categorical imperative would be one which presented an action as of itself objectively necessary, without regard to any other end.

Since every practical law presents a possible action as good and thus as necessary for a subject practically determinable by reason, all imperatives are formulas of the determination of action which is necessary by the principle of a will which is in any way good. If the action is good only as a means to something else, the imperative is *hypothetical*; but if it is thought of as good in itself, and hence as necessary in a will which of itself conforms to reason as the principle of this will, the imperative is *categorical*. [...]

#### [THE UNIVERSAL LAW FORMULA]

[...] There is, therefore, only one categorical imperative. It is: Act only according to that maxim by which you can at

the same time will that it should become a universal law. [...]

#### [ILLUSTRATIONS]

We shall now enumerate some duties, adopting the usual division of them into duties to ourselves and to others and into perfect and imperfect duties.

1. A man who is reduced to despair by a series of evils feels a weariness with life but is still in possession of his reason sufficiently to ask whether it would not be contrary to his duty to himself to take his own life. Now he asks whether the maxim of his action could become a universal law of nature. His maxim, however, is: For love of myself, I make it my principle to shorten my life when by a longer duration it threatens more evil than satisfaction. But it is questionable whether this principle of self-love could become a universal law of nature. One immediately sees a contradiction in a system of nature whose law would be to destroy life by the feeling whose special office is to impel the improvement of life. In this case it would not exist as nature; hence that maxim cannot obtain as a law of nature, and thus it wholly contradicts the supreme principle of all duty.

2. Another man finds himself forced by need to borrow money. He well knows that he will not be able to repay it, but he also sees that nothing will be loaned him if he does not firmly promise to repay it at a certain time. He desires to make such a promise, but he has enough conscience to ask himself whether it is not improper and opposed to duty to relieve his distress in such a way. Now, assuming he does decide to do so, the maxim of his action would be as follows: When I believe myself to be in need of money, I will borrow money and promise to repay it, although I know I shall never do so. Now this principle of self-love or of his own benefit may very well be compatible with his whole future welfare, but the question is whether it is right. He changes the pretension of self-love into a universal law and then puts the question: How would it be if my maxim became a universal law? He immediately sees that it could never hold as a universal law of nature and be consistent with itself; rather it must necessarily contradict itself. For the universality of a law which says that anyone who believes himself to be in need could promise what he pleased with the intention of not fulfilling it would make the promise itself and the end to be accomplished by it impossible; no one would believe what was promised to him but would only laugh at any such assertion as vain pretense.

3. A third finds in himself a talent which could, by means of some cultivation, make him in many respects a

useful man. But he finds himself in comfortable circumstances and prefers indulgence in pleasure to troubling himself with broadening and improving his fortunate natural gifts. Now, however, let him ask whether his maxim of neglecting his gifts, besides agreeing with his propensity to idle amusement, agrees also with what is called duty. He sees that a system of nature could indeed exist in accordance with such a law, even though man (like the inhabitants of the South Sea islands) should let his talents rust and resolve to devote his life merely to idleness, indulgence, and propagation — in a word, to pleasure. But he cannot possibly will that this should become a universal law of nature or that it should be implanted in us by a natural instinct. For, as a rational being, he necessarily wills that all his faculties should be developed, inasmuch as they are given to him for all sorts of possible purposes.

4. A fourth man, for whom things are going well, sees that others (whom he could help) have to struggle with great hardships, and he asks, “What concern of mine is it? Let each one be as happy as heaven wills, or as he can make himself; I will not take anything from him or even envy him; but to his welfare or to his assistance in time of need I have no desire to contribute.” If such a way of thinking were a universal law of nature, certainly the human race could exist, and without doubt even better than in a state where everyone talks of sympathy and good will, or even exerts himself occasionally to practice them while, on the other hand, he cheats when he can and betrays or otherwise violates the rights of man. Now although it is possible that

a universal law of nature according to that maxim could exist, it is nevertheless impossible to will that such a principle should hold everywhere as a law of nature. For a will which resolved this would conflict with itself, since instances can often arise in which he would need the love and sympathy of others, and in which he would have robbed himself, by such a law of nature springing from his own will, of all hope of the aid he desires.

#### [PERFECT AND IMPERFECT DUTIES]

The foregoing are a few of the many actual duties, or at least of duties we hold to be actual, whose derivation from the one stated principle is clear. We must be able to will that a maxim of our action become a universal law; this is the canon of the moral estimation of our action generally. Some actions are of such a nature that their maxim cannot even be *thought* as a universal law of nature without contradiction, far from it being possible that one could will that it should be such. In others this internal impossibility is not found, though it is still impossible to *will* that their maxim should be raised to the universality of a law of nature, because such a will would contradict itself. We easily see that the former maxim conflicts with the stricter or narrower (imprescriptible) duty, the latter with broader (meritorious) duty. Thus all duties, so far as the kind of obligation (not the object of their action) is concerned, have been completely exhibited by these examples in their dependence on the one principle. [...]