Groundwork of the Metaphysics of Morals

IMMANUEL KANT
1724–1804

Groundwork of the Metaphysics of Morals

Grundlegung zur Metaphysik der Sitten

Translated from the German with AI (Claude) and edited



FORAGE
forage.com

Groundwork of the Metaphysics of Morals · Immanuel Kant (1724–1804)

Translated from the German with AI (Claude) and edited. Reading edition, Forage, September 2026. forage.com/kant/groundwork

German text: the Akademie edition (AA vol. IV) via korpora.org (korpora.org/Kant/aa04), with emphasis from the Weischedel edition. English: translated from the German with AI (Claude), then edited for accuracy and consistency. Not reviewed by a human specialist. Free to reuse under CC BY-SA 4.0. The small numbers set in the text are the standard Akademie (AA) page references.

The original text is in the public domain. This translation is free to share and adapt under the Creative Commons Attribution-ShareAlike 4.0 International license (CC BY-SA 4.0): creativecommons.org/licenses/by-sa/4.0

Set in Source Serif 4.

CONTENTS

Introduction vii

Preface 1
Why a groundwork 6

Section One 9
The good will 10
The purpose of reason 11
Acting from duty 13
Duty and respect for the law 16
The false promise 19
Common reason and philosophy 21

Section Two 25
Against popular moral philosophy 28
The will and imperatives 32
Hypothetical and categorical imperatives 34
How imperatives are possible 37
The formula of universal law 41
The four examples 42
The need for an a priori principle 46
Humanity as an end in itself 49
The universally legislating will 53
The kingdom of ends 55
Dignity and price 57
Three formulas of one law 59
The autonomy of the will as the supreme principle of morality 63
The heteronomy of the will as the source of all spurious principles of morality 64
Classification of all possible principles of morality based on the assumed fundamental concept of heteronomy 65

Section Three 69
The concept of freedom is the key to the explanation of the autonomy of the will 69
Freedom must be presupposed as a property of the will of all rational beings 70
On the interest attaching to the ideas of morality 72
Two standpoints 74
How is a categorical imperative possible? 77
On the outermost boundary of all practical philosophy 79
Concluding remark 87

INTRODUCTION

In this short book (1785) Immanuel Kant set out to find "the supreme principle of morality." His answer is the categorical imperative: act only on a principle you could will to become a universal law.

Along the way come the ideas that still frame moral philosophy: that nothing is good without limitation except a good will; that an action has moral worth only when done from duty; that we must treat humanity never merely as a means but always also as an end; and that morality rests on the autonomy of a rational will.

The book has a preface and three sections, moving from ordinary moral thinking to philosophy and finally to the question of freedom.

Preface

Kant divides philosophy and calls for a pure moral philosophy, free of everything empirical.

387 Ancient Greek philosophy was divided into three sciences: *physics*, *ethics*, and *logic*. This division fits the nature of the subject perfectly, and nothing needs to be improved in it except perhaps to add the principle on which it rests. In that way we can make sure, on the one hand, that it is complete, and on the other, correctly determine the necessary subdivisions.

All rational cognition is either *material*, and considers some object, or *formal*, and is concerned merely with the form of the understanding and of reason itself, and with the universal rules of thinking in general, without distinction of objects. Formal philosophy is called *logic*. Material philosophy, however, which has to do with determinate objects and the laws to which they are subject, is in turn twofold. For these laws are either laws of *nature* or laws of *freedom*. The science of the former is called *physics*, that of the latter *ethics*; the former is also called the doctrine of nature, the latter the doctrine of morals.

Logic can have no empirical part—that is, no part in which the universal and necessary laws of thinking would rest on grounds taken from experience. For otherwise it would not be logic, that is, a canon for the understanding or for reason that holds for all thinking and must be demonstrated. Natural and moral philosophy, by contrast, can each have its empirical part. For natural philosophy must determine its laws for nature as an object of experience, and moral philosophy for the human will insofar as it is affected by nature: the former as laws according

to which everything [388] happens, the latter as laws according to which everything ought to happen, though also taking into account the conditions under which it often does not happen.

All philosophy insofar as it is based on grounds of experience can be called *empirical*; philosophy that sets forth its doctrines solely from principles *a priori* can be called *pure* philosophy. The latter, when it is merely formal, is called *logic*; but if it is restricted to determinate objects of the understanding, it is called *metaphysics*.

In this way there arises the idea of a twofold metaphysics: a *metaphysics of nature* and a *metaphysics of morals*. Physics will thus have its empirical part, but also a rational one; and so will ethics, although here the empirical part might be called, specifically, *practical anthropology*, while the rational part might properly be called *morals*.

All trades, crafts, and arts have gained from the division of labor. That is, no one makes everything; instead each person confines himself to a certain kind of work, clearly distinct from others in the way it must be handled, so that he can perform it with the greatest perfection and with greater ease. Where work is not differentiated and divided in this way, where everyone is a jack-of-all-trades, the trades still languish in the greatest barbarism. Now, it would in itself be a question not unworthy of consideration whether pure philosophy in all its parts does not require its own special practitioner. Would the learned trade as a whole not be better off if a warning were given to those who are accustomed to selling the empirical mixed with the rational, in all sorts of proportions unknown even to themselves, to suit the taste of the public—who call themselves independent thinkers, and call others, who prepare the merely rational part, hair-splitters? They would be warned not to carry on two occupations at once, which differ greatly in the way they must be handled, each of which perhaps requires a special talent, and whose combination in one person produces only bunglers. Still, here I ask only this: does not the nature of the science require that the empirical part always be carefully separated from the

rational part? Should we not put a metaphysics of nature before physics proper (empirical physics), and a metaphysics of morals before practical anthropology—each of which would have to be carefully cleansed of everything empirical? Only so can we know how much pure reason can [389] accomplish in both cases, and from what sources it draws this instruction of its own *a priori*. Whether this latter task is carried on by all teachers of morals (whose name is Legion) or only by some who feel a calling for it may be left open.[1]

Since my purpose here is really directed at moral philosophy, I restrict the question posed to this: do we not think it of the utmost necessity to work out, for once, a pure moral philosophy, completely cleansed of everything that may be merely empirical and belongs to anthropology? That there must be such a philosophy is self-evident from the common idea of duty and of moral laws. Everyone must admit that a law, if it is to hold morally, that is, as the ground of an obligation, must carry absolute necessity with it. Everyone must admit that the command "You shall not lie" does not hold merely for human beings, as though other rational beings need pay no heed to it—and likewise for all the other moral laws properly so called. Hence the ground of obligation here must not be sought in human nature, or in the circumstances of the world in which the human being is placed, but *a priori* solely in concepts of pure reason. And every other precept that is founded on principles of mere experience—even a precept that is universal in a certain respect—can indeed be called a practical rule, but never a moral law, insofar as it rests on empirical grounds in even the smallest part, perhaps only in its motive.

Thus moral laws, together with their principles, are not only essentially distinct from all other practical cognition in which there is anything empirical; rather, all moral philosophy rests entirely on its pure part. Applied to the human being, it does not borrow the least thing from knowledge of him (anthropology),

1 "Whose name is Legion": an allusion to Mark 5:9, where the possessed man gives his name as Legion, "for we are many."

but gives him, as a rational being, laws *a priori*. These laws, admittedly, still require a power of judgment sharpened by experience: partly to distinguish in which cases they apply, and partly to gain them access to the human will and force in being carried out. For the human being, affected as he is by so many inclinations, is indeed capable of the idea of a pure practical reason, but is not so easily able to make it effective *in concreto* [in the concrete] in the conduct of his life. A metaphysics of morals is therefore indispensably necessary. This is not merely from a motive of speculation, in order to investigate the source of the practical principles that lie in our reason *a* [390] *priori*, but because morals themselves remain subject to all sorts of corruption as long as that guiding thread and supreme norm for judging them correctly is lacking. For in the case of what is to be morally good, it is not enough that it *conform* to the moral law; it must also be done *for the sake of* that law. Otherwise the conformity is only very contingent and precarious, because an immoral ground will now and then produce actions that conform to the law, but more often actions contrary to it. Now the moral law in its purity and genuineness (which is precisely what matters most in practical affairs) is to be sought nowhere but in a pure philosophy. So this (metaphysics) must come first, and without it there can be no moral philosophy at all. Indeed, a philosophy that mixes those pure principles in with empirical ones does not even deserve the name of philosophy (for what distinguishes philosophy from common rational cognition is precisely that it presents in a separate science what common cognition grasps only as a mixture), much less the name of moral philosophy. For by this very mixing it even impairs the purity of morals themselves and works against its own end.

Let no one think that what is demanded here is already to be had in the propaedeutic that the celebrated *Wolff* placed before his moral philosophy—namely, in what he called *universal practical philosophy*—and that we are therefore not breaking entirely

new ground here.[2] Precisely because it was meant to be a universal practical philosophy, it did not consider a will of any special kind—say, one determined completely by principles *a priori*, without any empirical motives, which could be called a pure will—but volition in general, with all the actions and conditions that belong to it in this general sense. That is how it differs from a metaphysics of morals, just as general logic differs from transcendental philosophy. General logic sets forth the actions and rules of thinking *in general*; transcendental philosophy, merely the special actions and rules of *pure* thinking, that is, of the thinking through which objects are cognized completely *a priori*. For the metaphysics of morals is to investigate the idea and the principles of a possible *pure* will, and not the actions and conditions of human volition in general, which are for the most part drawn from psychology. That universal [391] practical philosophy also speaks (though without any warrant) of moral laws and of duty is no objection to my claim. For the authors of that science remain true to their idea of it in this respect as well. They do not distinguish the motives that are represented as such completely *a priori* by reason alone, and are properly moral, from the empirical motives that the understanding raises to universal concepts merely by comparing experiences. Instead, without paying attention to the difference in their sources, they consider them only according to their greater or smaller sum (since all of them are regarded as being of the same kind). In this way they form their concept of *obligation*. That concept is, to be sure, anything but moral; yet it is as good as can be demanded in a philosophy that makes no judgment at all about the *origin* of all possible practical concepts—whether they arise *a priori* or merely *a posteriori*.

2 Christian Wolff (1679–1754), the leading German rationalist philosopher of his day; his *Philosophia practica universalis* appeared in 1738–39.

WHY A GROUNDWORK

Kant explains why a pure moral philosophy is needed and outlines the book's three sections.

Intending some day to provide a metaphysics of morals, I first publish this groundwork. There is, to be sure, really no other foundation for such a metaphysics than a critique of *pure practical reason*, just as the foundation for metaphysics is the critique of pure speculative reason, which I have already published.[3] But, for one thing, the former critique is not as extremely necessary as the latter, because in moral matters human reason, even in the most common understanding, can easily be brought to a high degree of correctness and thoroughness, whereas in its theoretical but pure use it is wholly dialectical. For another, I require of a critique of pure practical reason, if it is to be complete, that it be possible at the same time to present its unity with speculative reason in a common principle; for in the end there can be only one and the same reason, which must differ merely in its application. But I could not yet bring the work to such completeness here without introducing considerations of an entirely different kind and confusing the reader. For this reason I have used the title *Groundwork of the Metaphysics of Morals* instead of *Critique of Pure Practical Reason*.

Third, however, a metaphysics of morals, despite its daunting title, is nevertheless capable of a high degree of popularity and suitability to the common understanding. I therefore find it useful to separate from it this preliminary work on its foundation, so that in future I need not attach to more accessible doctrines the subtleties [392] that are unavoidable here.

The present groundwork, however, is nothing more than the search for and establishment of *the supreme principle of morality*. This by itself constitutes a task complete in its purpose, and one to be kept separate from every other moral inquiry. My claims about this important main question, which until now has been

3 The *Critique of Pure Reason* (1781).

far from satisfactorily discussed, would no doubt gain much light from applying the same principle to the whole system, and great confirmation from the adequacy it displays everywhere. But I had to forgo this advantage, which at bottom would also be more self-serving than generally useful. For the ease with which a principle can be used, and its apparent adequacy, provide no entirely secure proof of its correctness. Rather, they arouse a certain bias against investigating and weighing it with full rigor in itself, without any regard to its consequences.

In this work I have adopted the method that I believe is most suitable if one wants to proceed analytically from common cognition to the determination of its supreme principle, and then in turn synthetically, from the examination of this principle and its sources back to common cognition, in which its use is found. The division has therefore turned out as follows:

1. *First Section*: Transition from common rational moral cognition to philosophical cognition.

2. *Second Section*: Transition from popular moral philosophy to the metaphysics of morals.

3. *Third Section*: Final step from the metaphysics of morals to the critique of pure practical reason.

Section One

Transition from common rational moral cognition to philosophical
Kant moves from the good will and duty to the principle of universal law.

[393] There is nothing anywhere in the world—indeed nothing even beyond it—that can possibly be thought of as good without limitation, except a *good will*. Understanding, wit, judgment, and whatever else the *talents* of the mind may be called, or courage, resoluteness, and perseverance in one's resolve, as qualities of *temperament*, are no doubt good and desirable in many respects. But they can also become extremely evil and harmful if the will that is to make use of these gifts of nature, and whose distinctive constitution is therefore called *character*, is not good. It is just the same with the *gifts of fortune*. Power, wealth, honor, even health and the complete well-being and contentment with one's condition that go by the name of *happiness*, produce boldness, and through it often arrogance too, unless a good will is present to correct their influence on the mind—and with it the whole principle of action—and make it serve universal ends. Not to mention that a rational and impartial spectator can never take any pleasure in seeing the uninterrupted prosperity of a being adorned by no trace of a pure and good will. Thus the good will seems to constitute the indispensable condition even of being worthy of happiness.

Some qualities are even conducive to this good will itself and can greatly ease its work. Nevertheless they have no [394] inner unconditional worth, but always presuppose a good will, which

limits the esteem that is otherwise rightly held for them and does not allow them to be considered absolutely good. Moderation in affects and passions, self-control, and sober reflection are not only good in many respects; they even seem to constitute part of the *inner* worth of a person. Yet they are far from deserving to be declared good without limitation (however unconditionally the ancients praised them). For without the principles of a good will they can become extremely evil; and the cold blood of a villain makes him not only far more dangerous, but also, immediately in our eyes, even more abhorrent than we would have taken him to be without it.

THE GOOD WILL

Kant argues that the good will is good in itself, whatever it achieves.

The good will is good not because of what it effects or accomplishes, not because of its fitness for attaining some proposed end, but solely because of its willing—that is, it is good in itself. Considered by itself, it is to be esteemed incomparably higher than anything that could ever be brought about by it in favor of some inclination, or indeed, if you like, of the sum of all inclinations. Suppose that, through some special disfavor of fate or the meager provision of a stepmotherly nature, this will wholly lacked the capacity to carry out its purpose. Suppose that even with its greatest effort it still accomplished nothing, and only the good will remained (not, to be sure, as a mere wish, but as the summoning of all means so far as they are in our power). Even then it would shine like a jewel for itself, as something that has its full worth in itself. Usefulness or fruitlessness can neither add anything to this worth nor take anything away from it. Usefulness would be, so to speak, only the setting, to make the jewel easier to handle in ordinary commerce, or to attract the attention of those who are not yet connoisseurs enough—but not to recommend it to connoisseurs or to determine its worth.

Yet in this idea of the absolute worth of the mere will, taking no account of any usefulness in estimating it, there is something

so strange that, despite all the agreement even of common reason with it, a suspicion must nevertheless arise. Perhaps it is secretly based on mere high-flown fantasy, and perhaps we have misunderstood nature's purpose in giving our will [395] reason as its ruler. Let us therefore examine this idea from this point of view.

THE PURPOSE OF REASON

Kant argues that reason's true purpose is not happiness but producing a will good in itself.

In the natural constitution of an organized being—that is, one purposively arranged for life—we take it as a principle that no instrument for any end will be found in it that is not also the most fitting for that end and the best suited to it. Now if, in a being that has reason and a will, its *preservation*, its *welfare*—in a word, its *happiness*—were the real end of nature, then nature would have made very poor arrangements for it in selecting the creature's reason to carry out this purpose. For all the actions the creature has to perform for this purpose, and the whole rule of its conduct, would have been marked out for it far more precisely by instinct, and that end could have been attained far more reliably in this way than it ever can be by reason. And if reason had been granted to the favored creature in addition, it would have had to serve only to let the creature reflect on the happy constitution of its nature, admire it, rejoice in it, and be grateful to the beneficent cause of it—not to subject its faculty of desire to that weak and deceptive guidance and meddle with nature's purpose. In a word, nature would have prevented reason from breaking out into *practical use* and from presuming, with its feeble insights, to devise for itself the plan of happiness and of the means to attain it. Nature would have taken over not only the choice of ends but also that of the means themselves, and with wise foresight would have entrusted both solely to instinct.

And in fact we find that the more a cultivated reason devotes itself to the purpose of enjoying life and happiness, the further a human being falls short of true contentment. From this there arises in many people—and indeed in those most practiced in the use of reason, if only they are candid enough to admit it—a certain degree of *misology*, that is, hatred of reason. For when they reckon up all the advantage they draw—I will not say from the invention of all the arts of common luxury, but even from the sciences (which in the end also seem to them a luxury of the understanding)—they nevertheless find that they have in fact only brought more hardship down [396] on their own heads than they have gained in happiness. And so in the end they envy, rather than look down on, the commoner run of people, who are closer to the guidance of mere natural instinct and do not allow their reason much influence on what they do and leave undone. And to this extent we must admit that the judgment of those who greatly tone down the boastful praises of the advantages that reason is supposed to procure for us with regard to happiness and contentment in life, and even reduce them below zero, is by no means sullen or ungrateful toward the goodness of the government of the world. Rather, these judgments are secretly based on the idea of another and far worthier purpose of our existence. It is to this purpose, and not to happiness, that reason is quite properly destined; and to it, as the supreme condition, the private purposes of the human being must for the most part give way.

For reason is not competent enough to guide the will reliably with regard to its objects and the satisfaction of all our needs (which it in part multiplies itself); an implanted natural instinct would have led to that end far more surely. Yet reason has nevertheless been imparted to us as a practical faculty, that is, as one that is to have influence on the *will*. So its true vocation must be to produce a will that is good, not perhaps *as a means* to some other purpose, but *good in itself*. For this, reason was absolutely necessary, if nature has everywhere proceeded purposively in distributing its capacities. This will, then, need not be the sole

and complete good, but it must be the highest good and the condition of everything else, even of all desire for happiness. In that case it is quite consistent with the wisdom of nature when we observe that the cultivation of reason, which is required for the first and unconditional purpose, in many ways limits—at least in this life—the attainment of the second, which is always conditional, namely happiness, and can even reduce it below nothing. Nature does not proceed unpurposively in this. For reason, which recognizes its highest practical vocation in the establishment of a good will, is capable, in attaining this purpose, only of a contentment of its own kind—namely, contentment from fulfilling an end that, again, only reason determines—even if this should involve considerable harm to the ends of inclination.

ACTING FROM DUTY

Kant distinguishes acting from duty from acting merely in conformity with duty or from inclination.

397 But we want to develop the concept of a will that is to be esteemed in itself and is good apart from any further purpose—a concept that already dwells in natural sound understanding and needs not so much to be taught as merely to be clarified; the concept that always ranks highest in estimating the whole worth of our actions, and constitutes the condition of all the rest. To do so, let us take up the concept of *duty*. It contains that of a good will, though under certain subjective limitations and hindrances. These, however, far from concealing it and making it unrecognizable, rather bring it out by contrast and make it shine forth all the more brightly.

I pass over here all actions that are already recognized as contrary to duty, even though they may be useful for this or that purpose; for with them the question whether they may have been done *from duty* does not even arise, since they actually conflict with it. I also set aside actions that really are in conformity with duty, but to which people immediately have *no inclination*, and which they nevertheless perform because they are driven to

them by some other inclination. For in that case it is easy to tell whether the action in conformity with duty was done *from duty* or for a self-seeking purpose. This distinction is far harder to notice where the action is in conformity with duty and the subject has, in addition, an *immediate* inclination to it. For example, it is certainly in conformity with duty that a shopkeeper not overcharge his inexperienced customer; and where there is a lot of trade, a prudent merchant does not do so either, but keeps a fixed general price for everyone, so that a child can buy from him just as well as anyone else. People are thus served *honestly*. But that is far from enough to make us believe that the merchant acted in this way from duty and principles of honesty; his advantage required it. It cannot be assumed here that he also had an immediate inclination toward his customers, so as, out of love as it were, to give none of them preference over another in price. So the action was done neither from duty nor from immediate inclination, but merely for a self-interested purpose.

On the other hand, to preserve one's life is a duty, and besides, everyone also has an immediate inclination to do so. But for that very reason the often anxious care that most people take of it has no inner worth, and its maxim has no moral 398
content. They preserve their life *in conformity with duty*, but not *from duty*. If, by contrast, adversity and hopeless grief have entirely taken away the taste for life; if the unfortunate man, strong of soul, more indignant at his fate than despondent or dejected, wishes for death and yet preserves his life without loving it—not from inclination or fear, but from duty—then his maxim has moral content.

To be beneficent where one can is a duty; and besides, there are many souls so sympathetically attuned that, even without any further motive of vanity or self-interest, they find an inner pleasure in spreading joy around them and can delight in the contentment of others insofar as it is their own work. But I maintain that in such a case an action of this kind, however much it conforms to duty and however amiable it may be, nevertheless has no true moral worth, but stands on a par with

other inclinations—for example, the inclination to honor. If this inclination fortunately lights upon what is in fact in the common interest and in conformity with duty, and hence honorable, it deserves praise and encouragement, but not esteem; for the maxim lacks moral content, namely that of doing such actions not from inclination but *from duty*. Suppose, then, that the mind of that friend of humanity were clouded by grief of his own, which extinguishes all sympathy with the fate of others. Suppose he still had the means to benefit others in need, but their distress did not move him, because he is sufficiently occupied with his own. And suppose that now, when no inclination any longer stirs him to it, he nevertheless tore himself out of this deadly insensibility and performed the action without any inclination, solely from duty. Then for the first time the action has its genuine moral worth. Further still: suppose nature had put little sympathy at all into the heart of this or that person. Suppose he (otherwise an honest man) were cold by temperament and indifferent to the sufferings of others—perhaps because, being himself equipped with the special gift of patience and enduring strength against his own sufferings, he presupposes the same in everyone else, or even demands it. Suppose nature had not really formed such a man (who would truly not be its worst product) to be a friend of humanity. Would he not still find within himself a source from which to give himself a far higher worth than that of a good-natured temperament may be? Certainly! It is precisely here that the worth of character begins, which is moral [399] and incomparably the highest: namely, that he is beneficent not from inclination but from duty.

To secure one's own happiness is a duty (at least indirectly); for lack of contentment with one's condition, under the pressure of many worries and amid unsatisfied needs, could easily become a great *temptation to transgress one's duties*. But even without looking to duty here, all human beings already have, of themselves, the most powerful and intimate inclination to happiness, because it is precisely in this idea that all inclinations are united in one sum. But the precept of happiness is for the

most part so constituted that it greatly impairs some inclinations, and yet a human being cannot form any determinate and reliable concept of the sum of the satisfaction of all of them under the name of happiness. It is therefore not surprising that a single inclination, determinate as to what it promises and the time within which it can be satisfied, can outweigh a wavering idea. A person suffering from gout, for example, can choose to enjoy what he likes and to suffer what he can, since by his reckoning he has at least not, on this occasion, deprived himself of the enjoyment of the present moment through perhaps groundless expectations of a happiness that is supposed to lie in health. But even in this case, if the universal inclination to happiness did not determine his will, and if health, at least for him, did not enter so necessarily into this reckoning, there still remains here, as in all other cases, a law: namely, to promote his happiness not from inclination but from duty. And only then does his conduct have its real moral worth.

It is undoubtedly in this way, too, that we should understand the passages of Scripture that command us to love our neighbor, even our enemy. For love as an inclination cannot be commanded; but beneficence from duty—even when no inclination at all drives us to it, and indeed when natural and unconquerable aversion resists it—is *practical* and not *pathological* love. It lies in the will and not in a propensity of feeling, in principles of action and not in melting sympathy; and only practical love can be commanded.[1]

DUTY AND RESPECT FOR THE LAW

Kant argues that moral worth lies in the maxim, and duty is acting from respect for the law.

The second proposition is this: an action from duty has its moral worth *not in the purpose* that is to be attained by it, but in the maxim according to which it is decided upon. It therefore does

1 See Leviticus 19:18 and Matthew 5:44. "Pathological" is Kant's term for what arises from feeling or sensibility (Greek *pathos*), not for anything diseased.

not depend on [400] the actuality of the object of the action, but merely on the *principle* of *volition* according to which the action has been done, without regard to any objects of the faculty of desire. That the purposes we may have in actions, and their effects, as ends and incentives of the will, cannot give actions any unconditional and moral worth is clear from what has gone before. In what, then, can this worth lie, if it is not to consist in the will in relation to the hoped-for effect of the actions? It can lie nowhere but *in the principle of the will*, without regard to the ends that can be brought about by such an action. For the will stands, as it were, at a crossroads, midway between its principle *a priori*, which is formal, and its incentive *a posteriori*, which is material. And since it must after all be determined by something, it will have to be determined by the formal principle of volition in general when an action is done from duty, since every material principle has been withdrawn from it.

The third proposition, as a consequence of the two previous ones, I would express as follows: *duty is the necessity of an action from respect for the law*. For an object as the effect of an action I have in view I can indeed have *inclination*, but *never respect*, precisely because it is merely an effect and not an activity of a will. In the same way, I cannot have respect for inclination as such, whether it is mine or someone else's. At most I can approve of it in the first case, and in the second I can sometimes even love it—that is, regard it as favorable to my own advantage. Only what is connected with my will merely as a ground and never as an effect—what does not serve my inclination but outweighs it, or at least excludes it entirely from the reckoning when a choice is made—only the bare law by itself, then, can be an object of respect and so a command. Now an action from duty is supposed to set aside entirely the influence of inclination, and with it every object of the will. So nothing is left that could determine the will except, objectively, the *law* and, subjectively, *pure respect*

for this practical law—and hence the maxim[2] of obeying such 401
a law even at the cost of all my inclinations.

The moral worth of an action therefore does not lie in the effect expected from it, and so not in any principle of action that needs to borrow its motive from this expected effect. For all these effects (the agreeableness of one's own condition, and even the promotion of other people's happiness) could also have been brought about by other causes, and so the will of a rational being was not needed for them. Yet it is in such a will alone that the highest and unconditional good can be found. Nothing but the *representation of the law* in itself—*which of course can occur only in a rational being*—can therefore constitute the pre-eminent good that we call moral, insofar as it is this representation, and not the hoped-for effect, that is the determining ground of the will. This good is already present in the person who acts on it; it need not first be awaited from the effect[3].

2 KANT'S NOTE. A *maxim* is the subjective principle of volition. The objective principle (that is, the one that would also serve all rational beings subjectively as a practical principle if reason had full control over the faculty of desire) is the practical *law*.

3 KANT'S NOTE. It might be objected that I am merely taking refuge in an obscure feeling behind the word *respect*, instead of giving a clear answer to the question through a concept of reason. But although respect is a feeling, it is not one *received* through outside influence. It is a feeling *self-produced* by a concept of reason, and so is specifically different from all feelings of the former kind, which can be traced back to inclination or fear. What I recognize immediately as a law for me, I recognize with respect, and respect signifies merely the consciousness of the *subordination* of my will to a law without the mediation of other influences on my senses. The immediate determination of the will by the law, and the consciousness of this, is called *respect*, so that respect is regarded as the *effect* of the law on the subject and not as its *cause*. Strictly speaking, respect is the representation of a worth that thwarts my self-love. It is therefore something regarded as an object neither of inclination nor of fear, although it has something analogous to both at once. The *object* of respect is thus the *law* alone—namely, the law that we impose *on ourselves* and yet as necessary in itself. As a law we are subject to it without consulting self-love; as imposed by ourselves, it is nonetheless a consequence of our will. In the first regard it is analogous to fear, in the second to inclination. All respect for a person is really only respect for the law (of integrity, and so on) of which that person gives us an example. Because we also regard the enlargement of our talents as a duty, we see in a person of talent, as it were, the *example of a law* (to become like that person in this by practice), and that is what constitutes our respect. All so-called moral *interest* consists solely in *respect* for the law.

[402] But what sort of law can it be, whose representation must determine the will even without regard to the effect expected from it, so that the will can be called good absolutely and without qualification? Since I have stripped the will of every impulse that could arise for it from obeying some particular law, nothing is left but the universal conformity of actions to law as such, and this alone is to serve the will as its principle. That is, I ought never to act except in such a way *that I could also will that my maxim should become a universal law*. Here, then, it is mere conformity to law as such (without taking as its basis any law determined for particular actions) that serves the will as its principle—and it must serve it so, if duty is not to be an empty delusion and a chimerical concept. Common human reason, in its practical judgments, agrees with this completely and always has this principle before its eyes.

THE FALSE PROMISE

Kant tests a lying promise and finds its maxim cannot be willed as a universal law.

Suppose, for example, the question is this: may I, when I am in a tight spot, make a promise with the intention of not keeping it? Here I easily distinguish the two meanings the question can have: whether it is prudent, or whether it is in conformity with duty, to make a false promise. The first may no doubt often be the case. I do see, of course, that it is not enough to get myself out of a present difficulty by means of this subterfuge. I must also consider carefully whether this lie might not later bring me much greater trouble than the troubles I am now escaping. And since, for all my supposed *cunning*, the consequences are not so easy to foresee, and trust once lost might prove far more damaging to me than all the harm I now mean to avoid, I must ask whether it would not be *more prudent* to proceed here on a universal maxim and make it my habit to promise nothing except with the intention of keeping it. But it soon becomes clear to me that such a maxim is still based only on the consequences I fear.

Now to be truthful from duty is something entirely different from being truthful out of fear of harmful consequences. In the first case the concept of the action already contains in itself a law for me; in the second I must first look around elsewhere to see what effects might be bound up with it for me. For if I deviate from the principle of duty, that is quite certainly evil; but if I desert my maxim of [403] prudence, that can sometimes be very advantageous to me, although it is of course safer to stick to it. Still, to teach myself in the shortest and yet most infallible way how to answer the question whether a lying promise is in conformity with duty, I ask myself: would I really be content for my maxim (to get myself out of a difficulty by an untruthful promise) to hold as a universal law, for myself as well as for others? And could I really say to myself: let everyone make an untruthful promise when he finds himself in a difficulty he cannot get out of in any other way? I then soon realize that I can indeed will the lie, but by no means a universal law to lie. For under such a law there would really be no promises at all. It would be pointless to declare my will regarding my future actions to others who would not believe this declaration—or who, if they rashly did, would pay me back in the same coin. Consequently my maxim, as soon as it was made a universal law, would have to destroy itself.

So I need no far-reaching ingenuity to know what I have to do in order that my volition be morally good. Inexperienced in the ways of the world, unable to be prepared for all the events that occur in it, I ask myself only: can you also will that your maxim become a universal law? If not, then it is to be rejected—and not because of some disadvantage it holds in store for you or for others, but because it cannot fit as a principle into a possible universal legislation. For this legislation, however, reason wrings from me immediate respect. I do not yet *see* what this respect is grounded on (that the philosopher may investigate), but I understand at least this much: that it is an estimation of a worth that far outweighs all the worth of whatever inclination recommends; and that the necessity of my actions from *pure* respect

for the practical law is what constitutes duty, to which every other motive must give way, because it is the condition of a will good *in itself*, whose worth surpasses everything.

COMMON REASON AND PHILOSOPHY

Kant argues that common reason already knows the moral principle, yet needs philosophy to resist its own doubts.

Thus, within the moral knowledge of common human reason, we have arrived at its principle. Common reason does not, of course, think this principle in abstraction, in such a universal form; yet it always actually has it before its eyes and uses it as the standard of its judgments. [404] It would be easy to show here how, with this compass in hand, it knows very well in every case that comes up how to distinguish what is good and what is evil, what conforms to duty and what is contrary to duty—if, without teaching it the least thing new, one merely makes it attend to its own principle, as Socrates did. It would be easy to show, too, that no science and no philosophy is needed to know what one has to do in order to be honest and good, and even to be wise and virtuous.[4] One might well have guessed in advance that knowledge of what every human being is obliged to do, and so also to know, would be the business of everyone, even the most common person. Here one cannot help admiring how far the practical faculty of judgment outstrips the theoretical in common human understanding. In theoretical matters, when common reason ventures to depart from the laws of experience and the perceptions of the senses, it falls into sheer incomprehensibilities and contradictions with itself—or at least into a chaos of uncertainty, obscurity and instability. In practical matters, by contrast, the power of judgment first begins to show itself to real advantage precisely when common understanding excludes all sensible incentives from practical laws. It then even becomes subtle, whether it wants to quibble with its conscience or with

4 Socrates (469–399 BC), the Athenian philosopher who, in Plato's dialogues, questions people until they recognize what they already implicitly know.

other claims about what is to be called right, or sincerely wants to determine the worth of actions for its own instruction. And, most remarkable of all, in the latter case it can hope to hit the mark just as well as any philosopher can promise himself—indeed, it is almost more reliable in this than the philosopher himself. For the philosopher has no principle other than the one common understanding has, but he can easily confuse his judgment with a mass of foreign considerations that do not belong to the matter, and make it swerve from the straight course. Would it not therefore be more advisable, in moral matters, to leave it at the judgment of common reason, and to bring in philosophy at most in order to present the system of morals more completely and comprehensibly, and its rules more conveniently for use (and still more for disputation)? But not in order to lead common human understanding, even for practical purposes, away from its happy simplicity and to set it, by means of philosophy, on a new path of inquiry and instruction?

Innocence is a glorious thing; but it is also, on the other hand,
405 very bad that it cannot protect itself well and is easily led astray. For this reason even wisdom—which otherwise consists more in doing and refraining than in knowing—still needs science as well, not in order to learn from it, but to gain acceptance and permanence for its precepts. The human being feels within himself a powerful counterweight to all the commands of duty, which reason presents to him as so deserving of the highest respect: namely, his needs and inclinations, whose entire satisfaction he sums up under the name of happiness. Now reason issues its precepts relentlessly, without promising anything to the inclinations, and so, as it were, slighting and disregarding those claims, which are so impetuous and at the same time seem so fair (and which refuse to be abolished by any command). From this, however, arises a *natural dialectic*: that is, a propensity to rationalize against those strict laws of duty, to cast doubt on their validity, or at least on their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations. That means corrupting them at their root and

robbing them of all their dignity—something that even common practical reason cannot, in the end, approve.[5]

So *common human reason* is driven to go beyond its own sphere and to take a step into the field of *practical philosophy*—not by any need of speculation (which never comes over it as long as it is content to be mere sound reason), but on practical grounds themselves. There it seeks information and clear direction about the source of its principle, and about the correct determination of this principle as against the maxims based on need and inclination. In this way it hopes to escape its predicament over the claims of both sides, and not to run the risk of losing all genuine moral principles through the ambiguity into which it easily falls. Thus, when practical common reason cultivates itself, a *dialectic* unfolds in it unnoticed, which compels it to seek help in philosophy, just as happens to it in its theoretical use. And so the former, no more than the latter, will find rest anywhere but in a complete critique of our reason.

5 "Dialectic" in Kant's sense: reason's tendency to entangle itself in plausible but illusory arguments.

Section Two

Transition from popular moral philosophy to the metaphysics of morals
Kant develops the categorical imperative in its formulas of universal law, humanity and autonomy.

406 Although we have so far drawn our concept of duty from the common use of our practical reason, it by no means follows that we have treated it as a concept of experience. On the contrary, when we attend to experience of what people do and leave undone, we meet frequent—and, as we ourselves admit, justified—complaints that one cannot cite a single sure example of the disposition to act from pure duty. Though much may happen *in conformity with* what *duty* commands, it always remains doubtful whether it really happens *from duty* and so has moral worth. Hence there have been philosophers in every age who have flatly denied that this disposition is actually present in human actions, and have ascribed everything to a more or less refined self-love. They did not, however, question the correctness of the concept of morality on that account. Rather, they spoke with heartfelt regret of the frailty and impurity of human nature: noble enough, they said, to make so respect-worthy an idea its precept, yet at the same time too weak to follow it, and using reason, which ought to serve it for giving laws, only to look after the interest of the inclinations—whether singly or, at best, in their greatest compatibility with one another.

407 In fact it is absolutely impossible to establish by experience, with complete certainty, a single case in which the maxim of an action otherwise in conformity with duty rested solely on

moral grounds and on the representation of one's duty. It is indeed sometimes the case that, with the keenest self-examination, we find nothing besides the moral ground of duty that could have been powerful enough to move us to this or that good action and to so great a sacrifice; but from this it cannot at all be inferred with certainty that no secret impulse of self-love, under the mere pretense of that idea, was actually the real determining cause of the will. For we like to flatter ourselves with a nobler motive that we falsely claim for ourselves; in fact, however, even by the most strenuous examination we can never get completely behind our secret incentives. For when moral worth is in question, what matters is not the actions one sees, but their inner principles, which one does not see.

Nor can one do those who laugh at all morality as a mere figment of human imagination, overreaching itself through conceit, a more welcome service than to concede that the concepts of duty must have been drawn solely from experience (just as, out of laziness, people like to persuade themselves that the same holds for all other concepts too). For that hands them a sure triumph. Out of love of humanity I will grant that most of our actions are still in conformity with duty. But if one looks more closely at all their scheming and striving, one everywhere comes upon the dear self, which always stands out; it is on this, and not on the strict command of duty, which would often demand self-denial, that their purpose rests. Nor does one need to be an enemy of virtue—only a cool observer who does not at once take the liveliest wish for the good for its reality—to become doubtful at certain moments (especially as one grows older, with a power of judgment partly made shrewder by experience and partly sharpened for observation) whether any true virtue is really to be found in the world at all. And here nothing can protect us from falling away entirely from our ideas of duty, or preserve in the soul a well-grounded respect for its law, except the clear conviction of the following. Even if there have never been actions that sprang from such pure sources, [408] the question here is not at all whether this or that happens; rather,

reason by itself, independently of all appearances, commands what ought to happen. Consequently, actions of which the world has perhaps so far given no example—whose very feasibility might be strongly doubted by someone who bases everything on experience—are nevertheless relentlessly commanded by reason. And so, for example, pure sincerity in friendship can be demanded of every human being no less, even if there may never yet have been a sincere friend, because this duty, as duty in general, lies prior to all experience in the idea of a reason that determines the will by *a priori* grounds.

Add to this that, unless one wants to deny the concept of morality all truth and all relation to any possible object, one cannot dispute that its law is of such wide significance that it must hold not merely for human beings but for all *rational beings as such*—not merely under contingent conditions and with exceptions, but with *absolute necessity*. It is then clear that no experience can give occasion to infer even the possibility of such apodictic laws.[1] For by what right can we bring into unlimited respect, as a universal precept for every rational nature, something that is perhaps valid only under the contingent conditions of humanity? And how are laws for the determination of *our* will to be taken as laws for the determination of the will of a rational being as such—and only as such for ours too—if they were merely empirical and did not have their origin completely *a priori* in pure but practical reason?

Nor could one give morality worse advice than to try to borrow it from examples. For every example of it that is presented to me must itself first be judged by principles of morality, to see whether it is worthy to serve as an original example, that is, as a model; it can by no means supply the concept of morality in the first place. Even the Holy One of the Gospel must first be compared with our ideal of moral perfection before he is recognized as such. He himself also says: Why do you call me (whom you see) good? No one is good (the archetype of the good) but the one God (whom you do not see). But where do we get the

1 "Apodictic": holding with strict necessity, not merely as a matter of fact.

concept of God [409] as the highest good?[2] Solely from the *idea* of moral perfection that reason sketches *a priori* and connects inseparably with the concept of a free will. Imitation has no place at all in morality, and examples serve only for encouragement. That is, they put beyond doubt the feasibility of what the law commands, and they make visible what the practical rule expresses more universally. But they can never entitle us to set aside their true original, which lies in reason, and to guide ourselves by examples.

AGAINST POPULAR MORAL PHILOSOPHY

Kant criticizes popular moral philosophy and insists that moral principles must be derived a priori.

If, then, there is no genuine supreme principle of morality that does not have to rest on pure reason alone, independently of all experience, I believe there is no need even to ask whether it is good to present these concepts in general (*in abstracto* [in the abstract]), as they stand firm *a priori* together with the principles belonging to them—provided that this knowledge is to be distinguished from common knowledge and called philosophical. But in our times it may well be necessary to ask. For if one took a vote on whether pure rational knowledge separated from everything empirical—and so a metaphysics of morals—or popular practical philosophy is to be preferred, one can soon guess which side would prevail.

This descent to popular concepts is certainly very commendable once the ascent to the principles of pure reason has first taken place and been completed to full satisfaction. That would mean first *grounding* the doctrine of morals on metaphysics and afterwards, once it stands firm, gaining it *acceptance* through popularity. But it is utterly absurd to want to cater to popularity already in the first inquiry, on which the whole correctness of the principles depends. Not only can this procedure never lay

2 Jesus' words in Mark 10:18 and Luke 18:19; the parentheses are Kant's own additions.

claim to the exceedingly rare merit of a true *philosophical popularity*, since there is no art at all in being generally intelligible if one thereby renounces all thorough insight. It also produces a revolting hodgepodge of patched-together observations and half-reasoned principles. Shallow minds feast on it, because it is something quite useful for everyday chatter; but people of insight feel confused by it and, dissatisfied yet unable to help themselves, turn their eyes away. Meanwhile philosophers, who see through the sham perfectly well, find little hearing when they call people away for a while from this supposed [410] popularity, so that they may rightly be popular only once definite insight has first been acquired.

One need only look at the essays on morality written in that favorite style. One will find now the particular vocation of human nature (but sometimes also the idea of a rational nature in general), now perfection, now happiness, here moral feeling, there fear of God, a little of this and a little of that, in a marvelous mixture. Yet it never occurs to anyone to ask whether the principles of morality are to be sought in knowledge of human nature at all (which we can have only from experience), and, if this is not so—if these principles are to be found completely *a priori*, free from everything empirical, solely in pure concepts of reason and nowhere else, not even in the smallest part—to adopt the plan of separating this inquiry off entirely as pure practical philosophy, or (if one may use so disreputable a name) as a metaphysics[3] of morals; of bringing it by itself to full completeness; and of asking the public, which demands popularity, to be patient until this undertaking is finished.

Such a completely isolated metaphysics of morals, however—mixed with no anthropology, no theology, no physics or hyper-

3 KANT'S NOTE. One can, if one likes, distinguish the pure philosophy of morals (metaphysics) from the applied (namely, applied to human nature), just as pure mathematics is distinguished from applied, and pure logic from applied. This terminology also reminds us at once that moral principles are not grounded on the peculiarities of human nature but must stand by themselves *a priori*, and that from such principles it must be possible to derive practical rules for every rational nature, and so for human nature too.

physics, and still less with occult qualities (which one might call hypophysical)—is not only an indispensable substrate of all theoretical, securely determined knowledge of duties. It is at the same time a desideratum of the highest importance for the actual fulfillment of their precepts.[4] For the pure representation of duty, and of the moral law in general, unmixed with any foreign addition of empirical enticements, has by way of reason alone (which here first becomes aware that it can also be practical by itself) an influence on the human heart so much more powerful than all other incentives[5] 411 that may be summoned from the empirical field, that reason, conscious of its dignity, despises those incentives and can gradually become their master. A mixed moral doctrine, by contrast, compounded of incentives from feelings and inclinations and at the same time of concepts of reason, must leave the mind wavering between motives that cannot be brought under any principle, and that can lead to the good only by chance, but often to evil as well.

From what has been said it is clear that all moral concepts have their seat and origin completely *a priori* in reason—and in the most common human reason just as much as in reason that is speculative in the highest degree; that they cannot be abstracted from any empirical and therefore merely contingent know-

4 "Hyperphysics" and "hypophysical": Kant's coinages for what lies above and below nature, respectively.

5 KANT'S NOTE. I have a letter from the late, excellent *Sulzer*, in which he asks me what the reason might be that the teachings of virtue, however much there is in them to convince reason, still accomplish so little. My answer was delayed because I was preparing to give it completely. But it is simply this: the teachers themselves have not clarified their concepts, and by trying to do too much—by drumming up motives for moral goodness on every side in order to make the medicine really strong—they spoil it. For the most common observation shows that when an act of integrity is presented as having been carried out with a steadfast soul, detached from every aim at any advantage in this world or another, even under the greatest temptations of need or allurement, it leaves far behind and eclipses every similar action that was affected even in the slightest by a foreign incentive. It elevates the soul and arouses the wish to be able to act like that too. Even children of middle years feel this impression, and duties should never be presented to them in any other way.

TRANSLATOR'S NOTE. Johann Georg Sulzer (1720–1779), Swiss-born philosopher and aesthetician in Berlin, who had written to Kant in 1770.

ledge; that it is precisely in this purity of their origin that their dignity lies, fitting them to serve us as supreme practical principles; that whenever one adds anything empirical, one subtracts just that much from their genuine influence and from the unlimited worth of actions; that it is not only a requirement of the greatest necessity for theoretical purposes, when speculation alone is at stake, but also of the greatest practical importance, to draw their concepts and laws from pure reason, to present them pure and unmixed, and indeed to determine the scope of this entire practical or pure rational knowledge— that is, the entire faculty of pure practical reason. In doing so, however, one must not make the principles dependent on the particular [412] nature of human reason, as speculative philosophy permits and indeed sometimes finds necessary. Rather, since moral laws are to hold for every rational being as such, one must derive them from the universal concept of a rational being as such. In this way all moral philosophy, which needs anthropology for its *application* to human beings, must first be presented independently of anthropology, as pure philosophy— that is, as metaphysics—and presented completely (which can well be done in this kind of wholly separated knowledge). And one must do this in the clear awareness that, without possessing such a metaphysics, it is not merely futile to determine precisely, for speculative judgment, the moral element of duty in everything that conforms to duty; it is impossible even in merely common and practical use, especially in moral instruction, to ground morals on their genuine principles, and thereby to produce pure moral dispositions and graft them onto people's minds for the highest good of the world.

In this work, however, we must advance by natural steps: not merely from common moral judgment (which is very worthy of respect here) to philosophical judgment, as has been done elsewhere, but from a popular philosophy, which goes no further than it can grope its way by means of examples, to metaphysics (which no longer lets itself be held back by anything empirical and, since it must survey the whole sum of rational knowledge

of this kind, goes if need be as far as ideas, where even examples desert us). To do this, we must follow the practical faculty of reason from its general rules of determination up to the point where the concept of duty springs from it, and present it clearly.

THE WILL AND IMPERATIVES

Kant defines the will as practical reason and explains imperatives as commands for imperfect wills.

Every thing in nature works according to laws. Only a rational being has the faculty of acting *according to the representation* of laws, that is, according to principles—in other words, has a *will*. Since *reason* is required for deriving actions from laws, the will is nothing other than practical reason. If reason infallibly determines the will, then the actions of such a being that are recognized 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, that is, as good. But if reason by itself alone does not sufficiently determine the will; if the will is also subject to subjective conditions (certain incentives) that do not always agree with the objective ones; in [413] a word, if the will is not *in itself* fully in accord with reason (as is actually the case with human beings): then actions that are cognized as objectively necessary are subjectively contingent, and the determination of such a will in accordance with objective laws is *necessitation*. That is, the relation of objective laws to a will that is not thoroughly good is represented as the determination of the will of a rational being by grounds of reason, to be sure, but by grounds that this will, by its nature, does not necessarily obey.

The representation of an objective principle, insofar as it necessitates a will, is called a command (of reason), and the formula of the 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 that, given its subjective constitution, is not necessarily determined by that

law (a necessitation). They say that it would be good to do or to refrain from something; but they say it to a will that does not always do a thing just because it is represented to it that the thing is good to do. What is practically *good*, however, is what determines the will by means of the representations of reason—hence not from subjective causes but objectively, that is, from grounds that are valid for every rational being as such. It is distinguished from the agreeable, which influences the will only by means of sensation, from merely subjective causes that hold only for this or that person's senses, and not as a principle of reason, which holds for everyone.[6]

414 A perfectly good will would therefore stand just as much under objective laws (of the good), but it could not be represented as *necessitated* by them to actions in conformity with law, because of itself, by its subjective constitution, it can be determined only by the representation of the good. Hence no imperatives hold for the *divine* will, or in general for a *holy* will; the *ought* is out of place here, because *willing* is already of itself necessarily in agreement with the law. Imperatives are therefore only formulas for expressing the relation of objective laws of willing in general to the subjective imperfection of the will of this or that rational being—for example, of the human will.

6 KANT'S NOTE. The dependence of the faculty of desire on sensations is called inclination, and so inclination always indicates a *need*. But the dependence of a contingently determinable will on principles of reason is called an *interest*. An interest is therefore found only in a dependent will, one that is not of itself always in accord with reason; in the divine will no interest can be conceived. But even the human will can *take an interest* in something without therefore *acting from interest*. The first signifies the *practical* interest in the action, the second the *pathological* interest in the object of the action. The first indicates only the dependence of the will on principles of reason in themselves; the second, its dependence on principles of reason for the sake of inclination, since here reason merely supplies the practical rule for how the need of inclination may be met. In the first case the action interests me; in the second, the object of the action (insofar as it is agreeable to me). We saw in the first section that in an action from duty one must look not to the interest in the object, but only to the interest in the action itself and in its principle in reason (the law).

HYPOTHETICAL AND CATEGORICAL IMPERATIVES

Kant distinguishes hypothetical imperatives of skill and prudence from the categorical imperative of morality.

Now all *imperatives* command either *hypothetically* or *categorically*. Hypothetical imperatives represent the practical necessity of a possible action as a means to attaining something else that one wills (or that it is at least possible one might will). The categorical imperative would be one that represented an action as objectively necessary for itself, without reference to any other end.

Every practical law represents a possible action as good, and therefore as necessary for a subject that is practically determinable by reason. So all imperatives are formulas for determining an action that is necessary according to the principle of a will that is good in some way. Now if the action would be good merely as a means *to something else*, the imperative is *hypothetical*; if the action is represented as *good in itself*, and hence as necessary in a will that in itself accords with reason, as that will's principle, then the imperative is *categorical*.

The imperative thus says which action possible through me would be good, and it represents the practical rule in relation to a will that does not straightaway perform an action just because it is good—partly because the subject does not always know that it is good, partly because, even if it did know this, its maxims could still be contrary to the objective principles of a practical reason.

The hypothetical imperative thus says only that the action is good for some *possible or actual* purpose. In the former [415] case it is a *problematically* practical principle, in the latter an *assertorically* practical one. The categorical imperative, which declares the action to be objectively necessary of itself without reference

to any purpose, that is, even without any other end, holds as an *apodictically* practical principle.[7]

Whatever is possible only through the powers of some rational being can also be thought of as a possible purpose for some will. Hence the principles of action, insofar as action is represented as necessary for achieving some possible purpose to be brought about by it, are in fact infinitely many. All sciences have some practical part, consisting of problems stating that some end is possible for us, and of imperatives stating how it can be attained. These can therefore in general be called imperatives of *skill*. Whether the end is rational and good is not at all the question here, but only what one must do to attain it. The prescriptions a physician follows to make his patient thoroughly healthy, and those a poisoner follows to kill his victim with certainty, are of equal worth to this extent: each serves to bring about its purpose perfectly. Since in early youth we do not know what ends may come our way in life, parents chiefly try to have their children learn a great *many* things, and they see to their *skill* in using means to all sorts of *arbitrary* ends. Of none of these ends can they determine whether it might actually become a purpose of their charge in the future, though it is always *possible* that he might one day have it. And this concern is so great that they commonly neglect to form and correct their children's judgment about the worth of the things they might choose to make their ends.

There is nonetheless *one* end that can be presupposed as actual in all rational beings (insofar as imperatives apply to them, namely as dependent beings), and thus one purpose that they not merely *can* have, but that one can safely presuppose they all *do* have by a necessity of nature: the purpose of *happiness*. The hypothetical imperative that represents the practical necessity of an action as a means to promoting happiness is *assertoric*. One may present it not merely as necessary for an uncertain, merely possible purpose, but for a purpose that one

7 Problematic, assertoric, apodictic: the modalities of judgment in Kant's logic—what is possibly, actually, or necessarily the case.

can presuppose, safely and *a priori*, in every human being, [416] because it belongs to his essence. Now skill in choosing the means to one's own greatest well-being can be called *prudence*[8] in the narrowest sense. So the imperative that refers to the choice of means to one's own happiness, that is, the precept of prudence, is still *hypothetical*: the action is commanded not absolutely, but only as a means to another purpose.[9]

Finally, there is an imperative that commands a certain conduct immediately, without presupposing as its condition any other purpose to be attained by that conduct. This imperative is *categorical*. It concerns not the matter of the action and what is to result from it, but the form and the principle from which the action itself follows; and what is essentially good in the action consists in the disposition, let the result be what it may. This imperative may be called the imperative *of morality*.

Willing according to these three kinds of principle is also clearly distinguished by the *dissimilarity* in the necessitation of the will. To make this noticeable as well, I believe they would be most suitably named, in their order, by saying that they are either *rules* of skill, or *counsels* of prudence, or *commands (laws)* of morality. For only *law* carries with it the concept of an *unconditional* and indeed objective and hence universally valid *necessity*; and commands are laws that must be obeyed, that is, followed even against inclination. *Counsel* does indeed involve necessity, but a necessity that can hold only under a subjective, contingent condition: whether this or that person counts this or that as part of his happiness. The categorical imperative, by contrast, is limited by no condition, and as absolutely—though practically—necessary it can quite properly be called a command.

8 KANT'S NOTE. The word prudence is taken in two senses: in the one it may bear the name worldly prudence, in the other that of private prudence. The first is a person's skill in influencing others so as to use them for his purposes. The second is the insight to unite all these purposes for his own lasting advantage. The latter is really the one to which even the worth of the former is traced back; and of someone who is prudent in the first way but not in the second, one might better say that he is clever and cunning, but on the whole imprudent.

9 "Prudence" (Klugheit) here means practical cleverness in securing one's own welfare, not a moral virtue.

One could also call the first kind of imperatives *technical* (belonging to art), the [417] second *pragmatic*[10] (belonging to welfare), and the third *moral* (belonging to free conduct in general, that is, to morals).

HOW IMPERATIVES ARE POSSIBLE

Kant asks how each kind of imperative is possible and finds the categorical one synthetic a priori.

Now the question arises: how are all these imperatives possible? This question does not ask how the performance of the action that the imperative commands can be conceived, but merely how the necessitation of the will that the imperative expresses in the task can be conceived. How an imperative of skill is possible surely needs no special discussion. Whoever wills the end also wills (so far as reason has decisive influence on his actions) the means to it that are indispensably necessary and in his power. As far as willing is concerned, this proposition is analytic. For in willing an object as my effect, my causality as acting cause, that is, the use of the means, is already thought; and the imperative draws the concept of actions necessary to this end out of the concept of willing this end. (Determining the means themselves to a proposed purpose certainly requires synthetic propositions; but these concern the ground not for making the act of will actual, but for making the object actual.) That in order to divide a line into two equal parts according to a secure principle I must draw two intersecting arcs from its ends—this, admittedly, mathematics teaches only through synthetic propositions. But that, if I know that the intended effect can come

10 KANT'S NOTE. It seems to me that the proper meaning of the word *pragmatic* can be most precisely determined in this way. For those *sanctions* are called pragmatic that flow, strictly speaking, not from the right of states as necessary laws, but from *provision* for the general welfare. A *history* is written pragmatically when it makes people *prudent*, that is, when it teaches the world how to look after its advantage better than, or at least as well as, earlier ages did.

TRANSLATOR'S NOTE. Pragmatic sanctions were solemn decrees of a ruler on affairs of state, such as the Pragmatic Sanction of 1713 regulating the Habsburg succession.

about only through such an action, then, if I fully will the effect, I also will the action required for it—this is an analytic proposition. For to represent something as an effect possible through me in a certain way, and to represent myself as acting in that same way with respect to it, are entirely the same.

If only it were as easy to give a determinate concept of happiness, the imperatives of prudence would agree entirely with those of skill and would be just as analytic. For here as there it would be said: whoever wills the end also wills (necessarily, in accordance with reason) the only means to it that are in [418] his power. Unfortunately, however, the concept of happiness is so indeterminate a concept that, although every human being wishes to attain happiness, he can never say determinately and consistently with himself what it is he really wishes and wills. The reason for this is that all the elements belonging to the concept of happiness are without exception empirical, that is, they must be borrowed from experience; and yet the idea of happiness requires an absolute whole, a maximum of well-being, in my present and every future condition. Now it is impossible for even the most insightful and at the same time most powerful, but still finite, being to form a determinate concept of what he actually wills here. If he wills wealth, how much care, envy and intrigue might he not bring down upon himself! If he wills a great deal of knowledge and insight, perhaps this would only give him a sharper eye, to show him all the more terribly the evils that are still hidden from him and yet cannot be avoided, or to load still more needs onto his desires, which already give him enough trouble. If he wills a long life, who will guarantee him that it would not be a long misery? If he wills at least health, how often has bodily discomfort kept someone from excesses into which unlimited health would have let him fall; and so on. In short, he is unable to determine with complete certainty, according to any principle, what would make him truly happy, because for that omniscience would be required. One cannot, therefore, act on determinate principles in order to be happy, but only on empirical counsels—for

example, of diet, thrift, courtesy, reserve and so on—which experience teaches promote well-being most on average. From this it follows that the imperatives of prudence, strictly speaking, cannot command at all, that is, cannot present actions objectively as *practically necessary*; that they are to be taken as counsels (*consilia*) rather than commands (*praecepta*) of reason; and that the task of determining, securely and universally, which action will promote the happiness of a rational being is completely insoluble. Hence no imperative is possible with respect to happiness that would, in the strict sense, command us to do what makes us happy. For happiness is an ideal not of reason but of imagination, resting merely on empirical grounds; and it is vain to [419] expect that these should determine an action by which the totality of a series of consequences that is in fact infinite would be attained. This imperative of prudence would nevertheless be an analytic practical proposition, if one assumes that the means to happiness could be specified with certainty. For it differs from the imperative of skill only in that for the latter the end is merely possible, while for the former it is given. Since both merely command the means to what one is presupposed to will as an end, the imperative that commands willing the means for whoever wills the end is analytic in both cases. So there is no difficulty either regarding the possibility of such an imperative.

By contrast, how the imperative of *morality* is possible is without doubt the only question in need of solution, since this imperative is not hypothetical at all, and so the objectively represented necessity cannot rest on any presupposition, as it does with hypothetical imperatives. But we must never lose sight of the fact that it cannot be settled *by any example*, and so empirically, whether there is any such imperative at all; rather, we must be concerned that all imperatives that seem categorical may nonetheless be covertly hypothetical. For example, suppose it is said: you ought not to make any deceitful promise; and suppose one assumes that the necessity of this omission is not mere counsel for avoiding some other evil—as if it said: you ought not

to make a lying promise, lest, if it comes out, you destroy your credit—but that an action of this kind must be regarded as evil in itself, so that the imperative of prohibition is categorical. Even so, one cannot show with certainty in any example that the will is determined here without any other incentive, merely by the law, even if it seems so. For it is always possible that fear of disgrace, or perhaps also an obscure apprehension of other dangers, may secretly influence the will. Who can prove by experience that a cause does not exist, when experience teaches nothing more than that we do not perceive it? In such a case, however, the so-called moral imperative, which as such appears categorical and unconditional, would in fact be only a pragmatic precept that draws our attention to our advantage and merely teaches us to take it into account.

We shall therefore have to investigate the possibility of a *categorical* imperative entirely *a priori*, since here we do not enjoy the [420] advantage that its reality is given in experience, so that its possibility would be needed not to establish it but merely to explain it. This much, however, can be seen in advance: that the categorical imperative alone has the tenor of a practical *law*, while all the others can indeed be called *principles* of the will, but not laws. For what is necessary to do merely to attain some arbitrary purpose can be regarded as in itself contingent, and we can always be released from the precept if we give up the purpose; whereas the unconditional command leaves the will no discretion to do the opposite, and so it alone carries with it the necessity that we require of a law.

Secondly, with this categorical imperative or law of morality, the ground of the difficulty (of seeing its possibility) is also very great. It is a synthetic practical proposition[11] *a priori*; and since

11 KANT'S NOTE. Without presupposing a condition drawn from any inclination, I connect the deed with the will *a priori*, hence necessarily (though only objectively, that is, under the idea of a reason that would have complete control over all subjective motives). This is therefore a practical proposition that does not derive the willing of an action analytically from some other willing already presupposed (for we have no such perfect will), but connects it immediately with the concept of the will of a rational being, as something not contained in that concept.

seeing the possibility of propositions of this kind is so difficult in theoretical cognition, it is easy to infer that it will be no less difficult in practical cognition.

With this task, we shall first try whether the mere concept of a categorical imperative might not also supply its formula, containing the proposition that alone can be a categorical imperative. For how such an absolute command is possible, even once we know how it runs, will still require special and difficult effort, which, however, we postpone to the last section.

When I think of a *hypothetical* imperative in general, I do not know in advance what it will contain until I am given the condition. But when I think of a *categorical* imperative, I know at once what it contains. For besides the law, the imperative contains only the necessity that the maxim[12] conform [421] to this law; but the law contains no condition to which it would be limited. So nothing is left but the universality of a law as such, to which the maxim of the action ought to conform; and it is this conformity alone that the imperative properly represents as necessary.

THE FORMULA OF UNIVERSAL LAW

Kant states the categorical imperative: act only on a maxim you can will as universal law.

There is therefore only a single categorical imperative, and it is this: *act only according to that maxim through which you can at the same time will that it become a universal law*.

Now if all imperatives of duty can be derived from this one imperative as from their principle, then—even though we leave it undecided whether what is called duty is not perhaps an empty concept altogether—we shall at least be able to show what we mean by it and what this concept is meant to say.

12 KANT'S NOTE. A *maxim* is the subjective principle of acting and must be distinguished from the *objective principle*, namely the practical law. The maxim contains the practical rule that reason determines in accordance with the conditions of the subject (often its ignorance, or else its inclinations), and so it is the principle on which the subject *acts*; the law, by contrast, is the objective principle, valid for every rational being, the principle on which it *ought to act*, that is, an imperative.

The universality of the law according to which effects take place constitutes what is properly called *nature* in the most general sense (as regards its form), that is, the existence of things insofar as it is determined by universal laws. So the universal imperative of duty could also run as follows: *act as if the maxim of your action were to become by your will a universal law of nature.*

THE FOUR EXAMPLES

Kant applies the formula to suicide, a false promise, neglected talents, and refusing to help others.

We shall now enumerate some duties, following their usual division into duties to ourselves and duties to other human beings, and into perfect and imperfect duties.[13]

1) Someone who, through a series of misfortunes that has grown to the point of hopelessness, feels weary of life is still 422 so far in possession of his reason that he can ask himself whether it might not be contrary to his duty to himself to take his own life. Now he tests whether the maxim of his action could become a universal law of nature. His maxim, however, is: from self-love I make it my principle to shorten my life if, by its longer duration, it threatens more ill than it promises pleasure. The only question left is whether this principle of self-love could become a universal law of nature. But one soon sees that a nature whose law it was to destroy life itself by means of the very feeling whose function is to impel the furtherance of life would contradict itself, and so would not subsist as nature. Hence that maxim cannot possibly hold as a universal law of

13 KANT'S NOTE. It should be carefully noted here that I reserve the division of duties entirely for a future *Metaphysics of Morals*; the division given here is therefore only an arbitrary one (to arrange my examples). For the rest, by a perfect duty I here understand one that permits no exception in favor of inclination; and so I have not merely external but also internal *perfect duties*. This runs contrary to the usage adopted in the schools, but I do not mean to defend it here, since for my purpose it makes no difference whether this is granted me or not.

TRANSLATOR'S NOTE. Kant published *The Metaphysics of Morals* in 1797.

nature, and consequently it wholly conflicts with the supreme principle of all duty.

2) Another finds himself driven by need to borrow money. He knows well that he will not be able to repay it, but he also sees that nothing will be lent to him unless he firmly promises to repay it at a set time. He would like to make such a promise; but he still has enough conscience to ask himself: is it not forbidden and contrary to duty to help oneself out of need in this way? Supposing he nonetheless decided to do it, his maxim of action would run: when I believe myself to be in need of money, I will borrow money and promise to repay it, even though I know this will never happen. Now this principle of self-love, or of one's own advantage, may perhaps be quite compatible with my whole future well-being; but the question now is whether it is right. I therefore turn the demand of self-love into a universal law and frame the question thus: how would things stand if my maxim became a universal law? I then see at once that it could never hold as a universal law of nature and agree with itself, but must necessarily contradict itself. For the universality of a law that everyone who believes himself to be in need could promise whatever occurs to him with the intention of not keeping it would make promising itself, and the end one may have in promising, impossible; for no one would believe that anything had been promised him, but would laugh at all such utterances as empty pretense.

3) A third finds in himself a talent that, with some [423] cultivation, could make him a person useful for all sorts of purposes. But he finds himself in comfortable circumstances and prefers to indulge in pleasure rather than trouble himself with enlarging and improving his fortunate natural predispositions. Yet he still asks whether his maxim of neglecting his natural gifts, besides agreeing of itself with his propensity to amusement, also agrees with what is called duty. He then sees that a nature could indeed still subsist under such a universal law, even if people (like the South Sea islanders) let their talents rust and were intent on devoting their lives merely to idleness, amuse-

ment and procreation—in a word, to enjoyment. But he cannot possibly *will* that this become a universal law of nature, or that it be implanted in us as such by natural instinct. For as a rational being he necessarily wills that all his capacities be developed, since they serve him and are given to him for all sorts of possible purposes.[14]

A *fourth*, who is doing well, while he sees that others must struggle with great hardships (whom he could easily help), thinks: what does it matter to me? Let each be as happy as Heaven wills or as he can make himself; I shall take nothing from him, nor even envy him; only I have no wish to contribute anything to his well-being or to his assistance in need! Now if such a way of thinking became a universal law of nature, the human race could admittedly subsist perfectly well—no doubt even better than when everyone chatters about sympathy and benevolence, and even takes pains to practice them on occasion, but also cheats wherever he can, sells out people's rights, or otherwise infringes them. But although it is possible that a universal law of nature could well subsist in accordance with that maxim, it is still impossible to *will* that such a principle hold everywhere as a law of nature. For a will that resolved on this would conflict with itself, since many cases can arise in which he needs the love and sympathy of others, and in which, by such a law of nature sprung from his own will, he would rob himself of all hope of the help he wishes for.

These, then, are some of the many actual duties, or at least of those we take to be such, whose classification on the basis of the one principle [424] cited above is clearly evident. We must *be able to will* that a maxim of our action become a universal law: this is the canon for the moral appraisal of our actions in general. Some actions are so constituted that their maxim cannot even be *thought* without contradiction as a universal law of nature, let alone that one could *will* that it *should* become one. In others this inner impossibility is not found, but it is still impossible to

14 An allusion to European travel reports of Pacific islanders, such as those from the voyages of Bougainville and Cook.

will that their maxim be raised to the universality of a law of nature, because such a will would contradict itself. It is easy to see that the first kind of action conflicts with strict or narrower (unremitting) duty, the second only with wider (meritorious) duty. And so, through these examples, all duties have been completely set out, as regards the kind of obligation (not the object of the action), in their dependence on the one principle.

Now if we attend to ourselves whenever we transgress a duty, we find that we do not actually will that our maxim should become a universal law, for that is impossible for us; rather, the opposite of our maxim should remain a universal law, only we take the liberty of making an *exception* to it for ourselves, or (even just this once) in favor of our inclination. Consequently, if we weighed everything from one and the same point of view, namely that of reason, we would find a contradiction in our own will: namely, that a certain principle be objectively necessary as a universal law, and yet subjectively not hold universally, but admit of exceptions. But since we consider our action at one time from the point of view of a will wholly in accord with reason, and then the very same action from the point of view of a will affected by inclination, there is actually no contradiction here, but rather a resistance of inclination to the precept of reason (*antagonismus* [antagonism]). Through it the universality of the principle (*universalitas* [universality]) is turned into mere general validity (*generalitas* [generality]), so that the practical principle of reason is to meet the maxim halfway. Although this cannot be justified in our own impartially rendered judgment, it nevertheless proves that we really acknowledge the validity of the categorical imperative and (with all respect for it) merely allow ourselves a few exceptions that seem to us insignificant and forced upon us.

THE NEED FOR AN A PRIORI PRINCIPLE

Kant argues that the categorical imperative must hold for all rational beings, grounded in reason alone.

425 We have thus shown at least this much: if duty is a concept that is to have meaning and to contain actual legislation for our actions, this can be expressed only in categorical imperatives and by no means in hypothetical ones. Likewise—and this is already a great deal—we have set out clearly, and determinately for every use, the content of the categorical imperative, which would have to contain the principle of all duty (if there were such a thing at all). But we have not yet got so far as to prove *a priori* that such an imperative actually exists: that there is a practical law which commands absolutely and of itself, without any incentives, and that obeying this law is duty.

With a view to getting there, it is of the utmost importance to take this as a warning: one must on no account let it enter one's head to try to derive the reality of this principle from the *particular property of human nature*. For duty is supposed to be the practical, unconditional necessity of action. It must therefore hold for all rational beings (the only beings to which an imperative can apply at all), and *only for that reason* be a law for every human will as well. By contrast, whatever is derived from the particular natural predisposition of humanity, from certain feelings and propensities, or even, if possible, from some particular tendency that would be peculiar to human reason and would not necessarily have to hold for the will of every rational being—that can indeed yield a maxim for us, but not a law. It can yield a subjective principle on which we have a propensity and inclination to act, but not an objective one on which we would be *directed* to act even if all our propensities, inclinations and natural make-up were against it. Indeed, the sublimity and inner dignity of the command in a duty is shown all the more, the less the subjective causes favor it and the more they oppose it—without

this in the least weakening the necessitation by the law or detracting anything from its validity.

Here, then, we see philosophy put in fact in a precarious position, one that is supposed to be firm even though it is neither suspended from anything in heaven nor supported by anything on earth. Here it is to prove its purity as the self-sustainer of its own laws, not as the herald of laws whispered to it by an implanted sense or by who knows what guardian nature. All of these, though they may always be better [426] than nothing at all, can never yield principles that reason dictates—principles that must have their source entirely and completely *a priori*, and with it their commanding authority: to expect nothing from human inclination but everything from the supremacy of the law and the respect owed to it, or, failing that, to condemn the human being to self-contempt and inner abhorrence.

Everything empirical, therefore, is not only wholly unfit to serve as an ingredient in the principle of morality; it is also highly detrimental to the purity of morals themselves. For the proper worth of an absolutely good will, a worth exalted above all price, consists precisely in this: that the principle of action is free from all influence of contingent grounds, which only experience can supply. Against this laxity, or even base way of thinking, which searches for the principle among empirical motives and laws, one cannot issue warnings too strongly or too often. For human reason, when weary, gladly rests on this pillow, and in a dream of sweet illusions (which make it embrace a cloud instead of Juno) it substitutes for morality a hybrid patched together from limbs of quite different lineage—one that looks

like whatever one wants to see in it, only not like virtue, for anyone who has once beheld virtue in her true form.[15][16]

The question, then, is this: is it a necessary law *for all rational beings* always to judge their actions by maxims that they themselves can will to serve as universal laws? If it is, then it must already be connected (completely *a priori*) with the concept of the will of a rational being as such. But to discover this connection one must, however much one resists, take a step beyond—namely into metaphysics, though into a region of it distinct from that of speculative philosophy: into the metaphysics of morals. [427] In a practical philosophy we are concerned not with assuming grounds of what *happens*, but with laws of what *ought to happen*, even if it never happens—that is, with objectively practical laws. There we need not investigate the grounds why something pleases or displeases; how the gratification of mere sensation differs from taste, and whether taste differs from a universal satisfaction of reason; what the feeling of pleasure and displeasure rests on, and how desires and inclinations arise from it, and from these, with the cooperation of reason, maxims. For all that belongs to an empirical psychology, which would make up the second part of the doctrine of nature, if this is regarded as *philosophy of nature* in so far as it is grounded on *empirical laws*. Here, however, we are speaking of objectively practical law, and hence of the relation of a will to itself in so far as it determines itself solely by reason. Everything that relates to the empirical then falls away of its own accord. For if *reason by itself alone* determines conduct (and it is the possibility of this that we now want to investigate), it must necessarily do so *a priori*.

The will is conceived as a capacity to determine itself to action *in accordance with the representation of certain laws*. And

15 KANT'S NOTE. To behold virtue in her proper form is nothing other than to present morality stripped of every admixture of the sensuous and of all the spurious adornment of reward or self-love. How far she then eclipses everything else that appears charming to the inclinations, anyone can easily discover by the slightest exercise of his reason, provided it has not been entirely spoiled for all abstraction.

such a capacity can be found only in rational beings. Now what serves the will as the objective ground of its self-determination is the *end*; and this, if it is given by mere reason, must hold equally for all rational beings. What, on the other hand, contains merely the ground of the possibility of an action whose effect is an end is called the *means*. The subjective ground of desiring is the *incentive*; the objective ground of willing is the *motive*. Hence the distinction between subjective ends, which rest on incentives, and objective ends, which depend on motives that hold for every rational being. Practical principles are *formal* when they abstract from all subjective ends; they are *material* when they are based on these, and hence on certain incentives. The ends that a rational being sets itself at its discretion as *effects* of its action (material ends) are without exception merely relative. For it is only their relation to a particularly constituted faculty of desire in the subject that gives them their worth. This worth therefore cannot furnish any universal principles—principles valid and necessary for all rational beings and also for every act of willing [428] —that is, practical laws. Hence all these relative ends are only the ground of hypothetical imperatives.

HUMANITY AS AN END IN ITSELF

Kant argues that rational beings are ends in themselves, never to be treated merely as means.

Suppose, however, that there were something *whose existence in itself* has an absolute worth, something that, as an *end in itself*, could be a ground of determinate laws. Then in it, and in it alone, would lie the ground of a possible categorical imperative, that is, of a practical law.

Now I say: the human being, and in general every rational being, *exists* as an end in itself, *not merely as a means* to be used at will by this or that will; in all its actions, whether directed toward itself or toward other rational beings, it must always be

16 Juno: in Greek myth Ixion, trying to seduce Hera (the Roman Juno), was given a cloud in her likeness to embrace.

regarded *at the same time as an end*. All objects of the inclinations have only a conditional worth; for if the inclinations, and the needs based on them, did not exist, their object would be without worth. But the inclinations themselves, as sources of need, are so far from having an absolute worth that would make them desirable in themselves that, on the contrary, to be entirely free of them must be the universal wish of every rational being. Thus the worth of any object *to be acquired* through our action is always conditional. Beings whose existence rests not on our will but on nature nevertheless have, if they are beings without reason, only a relative worth, as means, and are therefore called *things*. Rational beings, by contrast, are called *persons*, because their nature already marks them out as ends in themselves—that is, as something that may not be used merely as a means—and so to that extent limits all choice (and is an object of respect). These, then, are not merely subjective ends, whose existence as the effect of our action has a worth *for us*, but *objective ends*, that is, things whose existence is in itself an end—and indeed an end in whose place no other end can be put, one to which they would be meant to stand in service *merely* as means. For without this, nothing whatever of *absolute worth* would be found anywhere; and if all worth were conditional, and therefore contingent, then no supreme practical principle for reason could be found anywhere.

If, then, there is to be a supreme practical principle, and, with regard to the human will, a categorical imperative, it must be one that, from the representation of what is necessarily an end for everyone because it is an *end in itself*, constitutes an *objective* [429] principle of the will, and hence can serve as a universal practical law. The ground of this principle is: *rational nature exists as an end in itself*. The human being necessarily represents his own existence in this way; to that extent, then, it is a *subjective* principle of human actions. But every other rational being also represents its existence in this way, on the very same rational

ground that holds for me as well;[17] so it is at the same time an *objective* principle, from which, as a supreme practical ground, it must be possible to derive all laws of the will. The practical imperative will therefore be the following: *Act in such a way that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.* We shall see whether this can be carried out.

To keep to the previous examples: *First*, according to the concept of necessary duty to oneself, someone contemplating suicide will ask himself whether his action can be consistent with the idea of humanity *as an end in itself*. If he destroys himself in order to escape from a burdensome condition, he is making use of a person merely as *a means* to maintaining a tolerable condition until the end of life. But a human being is not a thing, and hence not something that can be used *merely* as a means; in all his actions he must always be regarded as an end in himself. So I cannot dispose of the human being in my own person by maiming, corrupting or killing him. (A more precise determination of this principle, to avoid any misunderstanding—for example, regarding the amputation of limbs to save myself, or exposing my life to danger in order to preserve my life, and so on—I must pass over here; it belongs to moral philosophy proper.)

Second, as regards necessary or owed duty to others: someone who has in mind to make a lying promise to others will see at once that he means to make use of another human being *merely as a means*, without the other at the same time containing the end in himself. For the person whom I want to use for my own purposes through such a promise cannot possibly agree to my way of treating him [430] and so cannot himself contain the end of this action. This conflict with the principle of other human beings is even more obvious if one brings in examples of attacks on the freedom and property of others. For then it is plain that someone who violates the rights of human beings intends to

17 KANT'S NOTE. I put forward this proposition here as a postulate. The grounds for it will be found in the last section.

make use of the person of others merely as a means, without considering that, as rational beings, they are always to be valued at the same time as ends—that is, only as beings who must also be able to contain in themselves the end of that very same action.[18]

Third, with regard to contingent (meritorious) duty to oneself, it is not enough that the action does not conflict with humanity in our person as an end in itself; it must also *harmonize with it*. Now there are in humanity predispositions to greater perfection, which belong to nature's end with regard to humanity in our subject. To neglect these might perhaps be consistent with the *preservation* of humanity as an end in itself, but not with the *advancement* of this end.

Fourth, concerning meritorious duty to others: the natural end that all human beings have is their own happiness. Now humanity could indeed subsist if no one contributed anything to the happiness of others, while also not intentionally taking anything away from it. But this is only a negative and not a positive agreement with *humanity as an end in itself*, if everyone did not also try, as far as he can, to advance the ends of others. For if that representation is to have its full effect in me, the ends of a subject who is an end in itself must, as far as possible, also be my ends.

This principle of humanity, and of every rational nature in general, *as an end in itself* (which is the supreme limiting [431] condition of the freedom of action of every human being), is not borrowed from experience. First, because of its universality: it

18 KANT'S NOTE. Let no one think that the trivial *quod tibi non vis fieri* etc. [what you do not want done to you, etc.] could serve here as a guideline or principle. For it is merely derived from the principle above, and with various limitations at that. It cannot be a universal law, since it contains the ground neither of duties to oneself, nor of duties of love to others (for many a person would gladly agree that others should not benefit him, if only he could be excused from showing them kindness), nor, finally, of owed duties to one another; for on this ground a criminal could argue against the judges who punish him, and so on.

TRANSLATOR'S NOTE. A Latin proverb, usually "Quod tibi fieri non vis, alteri ne feceris" ("Do not do to another what you do not want done to you"), a negative form of the Golden Rule.

applies to all rational beings as such, and no experience is sufficient to determine anything about them. Second, because in it humanity is represented not as an end of human beings (subjectively)—that is, as an object that one actually makes one's end of one's own accord—but as an objective end, which, whatever ends we may have, is to constitute as a law the supreme limiting condition of all subjective ends. Hence it must arise from pure reason.

THE UNIVERSALLY LEGISLATING WILL

Kant derives the idea of every rational will as universally legislating and names it autonomy.

For the ground of all practical legislation lies *objectively in the rule* and in the form of universality that makes it fit to be a law (possibly a law of nature), according to the first principle; but *subjectively* it lies in the *end*. The subject of all ends, however, is every rational being as an end in itself, according to the second principle. From this now follows the third practical principle of the will, as the supreme condition of its harmony with universal practical reason: the idea *of the will of every rational being as a universally legislating will*.

By this principle all maxims are rejected that cannot coexist with the will's own universal legislation. The will is thus not simply subject to the law, but subject to it in such a way that it must also be regarded as *legislating for itself*, and for that very reason as subject to the law in the first place (the law of which it can regard itself as the author).

The imperatives as previously represented—namely, as a lawfulness of actions universally resembling an *order of nature*, or as the universal *prerogative as ends* of rational beings in themselves—did indeed exclude from their commanding authority any admixture of an interest as incentive, precisely because they were represented as categorical. But they were only *assumed* to be categorical, because such an assumption had to be made if one wanted to explain the concept of duty. That there are prac-

tical propositions that command categorically could not itself be proved, just as it cannot be proved anywhere in this section either. But one thing could have been done: the renunciation of all interest in willing from duty, as the specific mark distinguishing the categorical from the hypothetical imperative, could have been indicated in the imperative itself, through [432] some determination it contained. And this is done in the present third formula of the principle, namely the idea of the will of every rational being as a *universally legislating will*.

For when we think of such a will, then although a will *that stands under laws* may still be bound to this law by means of an interest, a will that is itself supremely legislating cannot possibly, as such, depend on any interest. For such a dependent will would itself need yet another law, one that restricted the interest of its self-love to the condition of being valid as a universal law.

Thus the *principle* of every human will as *a will that legislates universally through all its maxims*,[19] provided it were otherwise correct, would be *well suited* to be the categorical imperative in this respect: that precisely for the sake of the idea of universal legislation it is *based on no interest*, and hence, of all possible imperatives, it alone can be *unconditional*. Or, better still, turning the proposition around: if there is a categorical imperative (that is, a law for every will of a rational being), it can only command that everything be done from the maxim of one's will as a will that could at the same time take itself, as universally legislating, for its object. For only then is the practical principle, and the imperative that the will obeys, unconditional, since it can have no interest whatever as its ground.

If we now look back on all the efforts ever undertaken to discover the principle of morality, it is no wonder that they all had to fail. People saw the human being bound to laws by his duty, but it did not occur to them that he is subject *only to his own* and yet *universal legislation*, and that he is obligated only to act in

19 KANT'S NOTE. I may be excused here from adducing examples to illustrate this principle, for those that first illustrated the categorical imperative and its formula can all serve the same purpose here.

accordance with his own will—a will, however, that according to nature's end is universally legislating. For if one thought of him only as subject to a law (whatever it might be), then this law had to carry with it some 433 interest, as attraction or constraint, because it did not arise as a law from *his* will; instead, his will was necessitated, in conformity with a law, by *something else* to act in a certain way. Through this entirely necessary conclusion, however, all the labor of finding a supreme ground of duty was irretrievably lost. For one never arrived at duty, but only at the necessity of an action from a certain interest. This might be one's own interest or someone else's. But then the imperative always had to turn out conditional, and could not serve at all as a moral command. I will therefore call this principle the principle of the *autonomy* of the will, in contrast with every other, which I accordingly count as *heteronomy*.[20]

THE KINGDOM OF ENDS

Kant describes the kingdom of ends, a union of rational beings under laws they give themselves.

The concept of every rational being as one that must regard itself as universally legislating through all the maxims of its will, in order to judge itself and its actions from this point of view, leads to a very fruitful concept that depends on it: namely, that *of a kingdom of ends*.

By a *kingdom*, however, I understand the systematic union of different rational beings through common laws. Now since laws determine ends according to their universal validity, then if we abstract from the personal differences between rational beings, and likewise from all the content of their private ends, we will be able to think of a whole of all ends in systematic connection—a whole both of rational beings as ends in themselves and of the

20 From the Greek: autonomy is "giving law to oneself"; heteronomy is "receiving law from another."

ends of their own that each may set itself—that is, a kingdom of ends, which is possible according to the above principles.[21]

For rational beings all stand under the *law* that each of them is to treat itself and all others *never merely as a means*, but always *at the same time as an end in itself*. From this, however, arises a systematic union of rational beings through common objective laws—that is, a kingdom. Since these laws have as their aim precisely the relation of these beings to one another as ends and means, it can be called a kingdom of ends (admittedly only an ideal).

A rational being belongs to the kingdom of ends as a *member* when, though it legislates universally in it, it is also itself subject to these laws. It belongs to it *as sovereign* when, as legislating, it is not subject to the will of any other.

434 A rational being must always regard itself as legislating in a kingdom of ends made possible by freedom of the will, whether as member or as sovereign. But it cannot claim the position of sovereign merely through the maxim of its will. It can do so only if it is a completely independent being, without needs and without any limitation of a capacity adequate to its will.

Morality thus consists in the relation of all action to the legislation through which alone a kingdom of ends is possible. This legislation, however, must be found in every rational being itself and must be able to arise from its will. The principle of this will is therefore: to do no action on any maxim except one that could also be consistent with its being a universal law, and hence only in such a way *that the will could regard itself through its maxim as at the same time universally legislating*. Now if maxims are not already by their nature necessarily in agreement with this objective principle of rational beings as universally legislating, then the necessity of acting according to that principle is called practical necessitation, that is, *duty*. Duty does not apply to the sovereign in the kingdom of ends, but it does apply to every member, and indeed to all in equal measure.

21 Kant's *Reich* can also mean "realm"; "kingdom" keeps the image of a community united under common laws.

The practical necessity of acting according to this principle, that is, duty, does not rest at all on feelings, impulses and inclinations, but solely on the relation of rational beings to one another—a relation in which the will of a rational being must always be regarded at the same time as *legislating*, since otherwise it could not think of them as an *end in itself*. Reason thus relates every maxim of the will, as universally legislating, to every other will and also to every action toward oneself. It does so not for the sake of any other practical motive or future advantage, but from the idea of the *dignity* of a rational being that obeys no law except one that it at the same time gives itself.

DIGNITY AND PRICE

Kant distinguishes things with a price from persons, whose dignity rests on autonomy.

In the kingdom of ends everything has either a *price* or a *dignity*. Whatever has a price can be replaced by something else as its *equivalent*; whatever, on the other hand, is exalted above all price, and so admits of no equivalent, has a dignity.

What relates to general human inclinations and needs has a *market price*. What, even without presupposing a need, accords with a certain taste—that is, with a satisfaction in the 435 mere purposeless play of our mental powers—has a *fancy price*. But that which constitutes the condition under which alone something can be an end in itself has not merely a relative worth, that is, a price, but an inner worth, that is, *dignity*.[22]

Now morality is the condition under which alone a rational being can be an end in itself, since only through it is it possible to be a legislating member in the kingdom of ends. Hence morality, and humanity in so far as it is capable of morality, is that which alone has dignity. Skill and diligence in work have a market price; wit, lively imagination and good humor have a fancy price; but fidelity in promising and benevolence from principle

22 "Fancy price" renders Kant's *Affectionspreis*: the value a thing has because it pleases our taste or fancy, not because it meets a need.

(not from instinct) have an inner worth. Neither nature nor art contains anything that could take their place if they were lacking. For their worth consists not in the effects that arise from them, nor in the advantage and utility they produce, but in the dispositions, that is, the maxims of the will, which are ready to reveal themselves in this way in actions, even if success does not favor them. Nor do these actions need any recommendation from some subjective bent or taste to make us regard them with immediate favor and satisfaction; they need no immediate propensity or feeling for them. They present the will that performs them as the object of an immediate respect, and nothing but reason is required to *impose* them on the will—not to *coax* them out of it, which in the case of duties would in any case be a contradiction. This estimation therefore reveals the worth of such a way of thinking as dignity, and places it infinitely above all price, with which it cannot be weighed or compared at all without, as it were, violating its sanctity.

And what is it, then, that entitles the morally good disposition, or virtue, to make such lofty claims? It is nothing less than the *share* it gives a rational being *in universal legislation*, which makes it fit to be a member of a possible kingdom of ends. For this it was already destined by its own nature: as an end in itself, and for that very reason as legislating in the kingdom of ends—free with respect to all laws of nature, obeying only those that it gives itself and according to which its maxims can belong to a universal legislation (to which it at the same time [436] subjects itself). For nothing has any worth except the worth that the law determines for it. But the legislation itself, which determines all worth, must for that very reason have a dignity, that is, an unconditional, incomparable worth; and the word *respect* alone provides a fitting expression for the esteem that a rational being must have for it. *Autonomy* is therefore the ground of the dignity of human nature and of every rational nature.

THREE FORMULAS OF ONE LAW

Kant presents the three formulas as versions of one law and summarizes the argument so far.

The three ways of representing the principle of morality set out above are, however, at bottom only so many formulas of the very same law, any one of which of itself unites the other two within it. There is nevertheless a difference among them, though it is subjectively rather than objectively practical: it serves to bring an idea of reason closer to intuition (by a certain analogy), and thereby closer to feeling. For all maxims have

1) a *form*, which consists in universality; and here the formula of the moral imperative is expressed thus: maxims must be chosen as if they were to hold as universal laws of nature;

2) a *matter*, namely an end; and here the formula says: the rational being, as an end by its nature and hence as an end in itself, must serve for every maxim as the limiting condition of all merely relative and arbitrary ends;

3) *a complete determination* of all maxims through that formula, namely: all maxims, from one's own legislation, are to harmonize into a possible kingdom of ends, as a kingdom of nature.[23] The progression here takes place as through the categories of the *unity* of the form of the will (its universality), the *plurality* of the matter (the objects, that is, the ends), and the *allness* or totality of the system of these. But in moral *judging* one does better always to proceed by the strict method and to take as one's basis the universal formula of the categorical imperative: *act according to the maxim that can at the same time make itself a universal law* 437 .[24] But if one wants at the same time to gain *entry*

23 KANT'S NOTE. Teleology considers nature as a kingdom of ends; morals considers a possible kingdom of ends as a kingdom of nature. In the former, the kingdom of ends is a theoretical idea for explaining what exists. In the latter, it is a practical idea for bringing about what does not exist but can become actual through what we do and refrain from doing—and indeed for bringing it about in accordance with this very idea.

24 Unity, plurality and totality (allness) are the three categories of quantity in Kant's *Critique of Pure Reason*.

for the moral law, it is very useful to lead one and the same action through the three concepts named, and thereby, as far as possible, to bring it closer to intuition.

We can now end where we set out at the beginning, namely with the concept of an unconditionally good will. *That will is absolutely good* which cannot be evil—hence whose maxim, when made into a universal law, can never conflict with itself. This principle is therefore also its supreme law: act always according to that maxim whose universality as a law you can at the same time will. This is the only condition under which a will can never be in conflict with itself, and such an imperative is categorical. Since the validity of the will as a universal law for possible actions is analogous to the universal connection of the existence of things according to universal laws, which is the formal aspect of nature in general, the categorical imperative can also be expressed thus: *act according to maxims that can at the same time have themselves as universal laws of nature for their object*. Such, then, is the formula of an absolutely good will.

Rational nature distinguishes itself from the rest of nature by setting itself an end. This end would be the matter of every good will. But in the idea of a will that is absolutely good without any limiting condition (the attainment of this or that end), we must abstract entirely from every end to be *brought about* (as something that would make every will only relatively good). So the end must here be thought of not as an end to be brought about, *but as a self-standing* end, and hence only negatively: that is, as an end that must never be acted against, and that must therefore never be valued merely as a means but always at the same time as an end in every act of willing. Now this end can be nothing other than the subject of all possible ends itself, because this subject is at the same time the subject of a possible absolutely good will; for such a will cannot without contradiction be ranked below any other object. The principle "Act in relation to every rational being (yourself and others) in such a way that it counts in your maxim at the same time as an end in itself" is therefore at bottom the same as the principle "Act according to a

maxim that at the same time contains within itself its own universal validity for every rational being." [438] For to say that, in using means to any end, I ought to restrict my maxim to the condition of its universal validity as a law for every subject amounts to saying this: the subject of ends, that is, the rational being itself, must be made the basis of all maxims of actions never merely as a means, but as the supreme limiting condition in the use of all means, that is, always at the same time as an end.

Now it follows indisputably from this that every rational being, as an end in itself, must be able to regard itself, with respect to any law whatever to which it may be subject, as at the same time universally legislating. For it is precisely this fitness of its maxims for universal legislation that marks it out as an end in itself. It follows likewise that its dignity (prerogative) over all mere beings of nature brings with it the need always to take its maxims from the point of view of itself, and at the same time from that of every other rational being as a legislating being (such beings are for that reason also called persons). Now in this way a world of rational beings (*mundus intelligibilis* [intelligible world]) is possible as a kingdom of ends—possible, namely, through the legislation of all persons themselves as its members. Accordingly, every rational being must act as if it were, through its maxims, always a legislating member in the universal kingdom of ends. The formal principle of these maxims is: act as if your maxim were at the same time to serve as a universal law (for all rational beings). A kingdom of ends is thus possible only by analogy with a kingdom of nature. But the former is possible only through maxims, that is, rules imposed on oneself; the latter only through laws of efficient causes necessitated from outside. Even so, the whole of nature, although it is regarded as a machine, is still given the name of a kingdom of nature, on the ground that it relates to rational beings as its ends. Now such a kingdom of ends would actually come into being through maxims whose rule the categorical imperative prescribes to all rational beings, *if they were universally followed*. Yet a rational being cannot count on every other being keeping

faith with this maxim just because it itself follows it punctiliously. Nor can it count on the kingdom of nature and its purposive order harmonizing with it, as a fitting member, toward a kingdom of ends made possible through itself—that is, on nature favoring its expectation of happiness [439] . Even so, that law—act according to the maxims of a universally legislating member of a merely possible kingdom of ends—remains in full force, because it commands categorically. And here lies precisely the paradox: that the mere dignity of humanity as rational nature, without any further end or advantage to be gained by it—hence respect for a mere idea—should nonetheless serve as an inflexible precept of the will; and that the sublimity of the maxim consists precisely in this independence from all such incentives, as does the worthiness of every rational subject to be a legislating member in the kingdom of ends. For otherwise the subject would have to be represented only as subject to the natural law of its needs. Even if the kingdom of nature and the kingdom of ends were thought of as united under one head, so that the latter no longer remained a mere idea but obtained true reality, this would indeed give that idea the addition of a strong incentive, but never any increase in its inner worth. For even so, this sole, unlimited legislator would still always have to be represented as judging the worth of rational beings only by their disinterested conduct, which they prescribe to themselves merely on the basis of that idea. The essence of things is not changed by their external relations; and whatever, without regard to such relations, alone constitutes the absolute worth of a human being is what he must be judged by, whoever the judge may be—even the supreme being. *Morality* is therefore the relation of actions to the autonomy of the will, that is, to a possible universal legislation through its maxims. An action that is compatible with the autonomy of the will is *permitted*; one that does not agree with it is *forbidden*. A will whose maxims necessarily harmonize with the laws of autonomy is a *holy*, absolutely good will. The dependence of a will that is not absolutely good on the principle of autonomy (moral necessitation) is *obligation*. Obligation,

then, cannot apply to a holy being. The objective necessity of an action from obligation is called *duty*.

From what has just been said, it is now easy to explain how it comes about that, although under the concept of duty we think of subjection to the law, we nonetheless at the same time [440] represent a certain sublimity and *dignity* in the person who fulfills all his duties. For there is no sublimity in the person insofar as he is *subject* to the moral law; but there is insofar as he is at the same time *legislating* with respect to that very law, and subordinate to it only for that reason. We have also shown above how neither fear nor inclination, but solely respect for the law, is the incentive that can give an action moral worth. Our own will, insofar as it would act only under the condition of a universal legislation possible through its maxims—this will, possible for us in idea, is the proper object of respect. And the dignity of humanity consists precisely in this capacity to legislate universally, though on condition of being at the same time itself subject to this very legislation.

THE AUTONOMY OF THE WILL AS THE SUPREME PRINCIPLE OF MORALITY

Kant defines autonomy, the will being a law to itself, as the sole principle of morality.

Autonomy of the will is the property of the will by which it is a law to itself (independently of any property of the objects of volition). The principle of autonomy is therefore: to choose only in such a way that the maxims of one's choice are at the same time included as universal law in the same volition. That this practical rule is an imperative—that is, that the will of every rational being is necessarily bound to it as a condition—cannot be proved by mere analysis of the concepts found in it, because it is a synthetic proposition. One would have to go beyond the cognition of objects to a critique of the subject, that is, of pure practical reason; for this synthetic proposition, which commands apodictically, must be capable of being cognized entirely *a priori*. But that task does not belong in the present section.

That the principle of autonomy in question is the sole principle of morals, however, can very well be shown by mere analysis of the concepts of morality. For such analysis reveals that its principle must be a categorical imperative, and that this imperative commands neither more nor less than precisely this autonomy.

THE HETERONOMY OF THE WILL AS THE SOURCE OF ALL SPURIOUS PRINCIPLES OF MORALITY

Kant argues that seeking the moral law in any object of the will yields only hypothetical imperatives.

441 If the will seeks the law that is to determine it *anywhere else* than in the fitness of its maxims for its own universal legislation—if, therefore, going beyond itself, it seeks this law in the character of any of its objects—the result is always *heteronomy*. The will then does not give itself the law; instead, the object, through its relation to the will, gives the law to it. This relation, whether it rests on inclination or on representations of reason, makes only hypothetical imperatives possible: I ought to do something *because I want something else*. The moral, and hence categorical, imperative, by contrast, says: I ought to act in this way or that even if I wanted nothing else. For example, the former says: I ought not to lie if I want to keep my good name; the latter says: I ought not to lie even if lying would not bring me the slightest disgrace. The latter imperative must therefore abstract from every object so far that the object has no *influence* at all on the will, so that practical reason (will) does not merely administer an interest not its own but simply demonstrates its own commanding authority as supreme legislation. Thus I ought, for example, to try to promote the happiness of others, not as though its existence mattered anything to me (whether through immediate inclination or through some satisfaction gained indirectly through reason), but merely because a maxim that excludes it cannot be comprehended as a universal law in one and the same volition.

CLASSIFICATION OF ALL POSSIBLE PRINCIPLES OF MORALITY BASED ON THE ASSUMED FUNDAMENTAL CONCEPT OF HETERONOMY

Kant rejects moral theories based on happiness, moral feeling, perfection and the will of God.

Here, as everywhere in its pure use, human reason, so long as it lacks a critique, first tries every possible wrong path before it succeeds in hitting on the one true path. All principles that may be adopted from this point of view are either *empirical or rational*. The *former*, drawn from the principle [442] of *happiness*, are built on physical or moral feeling; the *latter*, drawn from the principle of *perfection*, are built either on the rational concept of perfection as a possible effect, or on the concept of an independent perfection (the will of God) as the determining cause of our will. *Empirical principles* are altogether unfit to serve as a foundation for moral laws. For the universality with which such laws are to hold for all rational beings without distinction—the unconditional practical necessity thereby imposed on them—falls away if their ground is taken from the *particular constitution of human nature*, or from the contingent circumstances in which it is placed. The principle of *one's own happiness*, however, is the most objectionable. This is not merely because it is false, and experience contradicts the pretense that well-being always matches good conduct. Nor is it merely because it contributes nothing whatever to establishing morality, since making someone happy is quite different from making him good, and making him prudent and sharp about his own advantage is quite different from making him virtuous. It is rather because it supplies morality with incentives that undermine it and destroy all its sublimity: they put the motives to virtue in the same class as the motives to vice and teach only how to calculate better, while wiping out the specific difference between the two entirely;

whereas moral feeling, this supposed special sense[25], still stays closer to morality and its dignity. (Appealing to it is shallow indeed: those who cannot *think* believe they can help themselves out by *feeling*, even in matters that depend solely on universal laws; feelings, which by nature differ infinitely from one another in degree, provide no uniform standard of good and evil; and no one can judge validly for others by means of his own feeling.) It stays closer because it does virtue the honor of ascribing the satisfaction and esteem we feel for it to virtue *immediately*, and does not, as it were, [443] tell virtue to its face that what binds us to it is not its beauty but only our advantage.

Among the *rational* grounds of morality—those based on reason—the ontological concept of *perfection* (however empty and indeterminate it is, and hence useless for finding, in the immeasurable field of possible reality, the greatest sum suitable for us; and however much, in trying to distinguish the reality under discussion here specifically from every other, it has an unavoidable tendency to go round in a circle and cannot avoid covertly presupposing the morality it is supposed to explain) is nevertheless better than the theological concept, which derives morality from a divine, supremely perfect will. This is not merely because we cannot intuit God's perfection and can only derive it from our own concepts, the foremost of which is that of morality. It is rather because, if we do not do this (and doing it would be a crude circle in explanation), the only concept of his will left to us—made up of the attributes of lust for glory and dominion, combined with fearful representations of power and vindictiveness—would have to be the foundation for a system of morals directly opposed to morality.

25 KANT'S NOTE. I count the principle of moral feeling as part of the principle of happiness, because every empirical interest promises a contribution to our well-being through the agreeableness that something affords—whether this happens immediately and without a view to advantage, or with regard to it. Likewise, one must, with *Hutcheson*, count the principle of sympathy with the happiness of others as part of the same moral sense that he assumed.

TRANSLATOR'S NOTE. Francis Hutcheson (1694–1746), Irish-Scottish philosopher who held that we judge right and wrong through an innate "moral sense."

But if I had to choose between the concept of a moral sense and that of perfection in general (neither of which at least does harm to morality, although they are of no use at all in supporting it as its foundations), I would decide for the latter. For since it at least withdraws the decision of the question from sensibility and brings it before the court of pure reason, it preserves the indeterminate idea (of a will good in itself) unfalsified for more precise determination, even though here too it decides nothing.

For the rest, I believe I may be excused from an extended refutation of all these doctrines. Refuting them is so easy, and is presumably so well understood even by those whose office requires them to declare for one of these theories after all (because audiences do not readily tolerate a suspension of judgment), that it would only be wasted labor. What interests us more here, however, is to know this: that these principles never set up anything but heteronomy of the will as the first ground of morality, and for that very reason must necessarily fail to achieve their end.

444 Wherever an object of the will has to be laid down as the basis for prescribing to the will the rule that determines it, the rule is nothing but heteronomy. The imperative is conditioned, namely: *if* or *because* one wills this object, one ought to act in this way or that. Hence it can never command morally, that is, categorically. Whether the object determines the will by means of inclination, as in the principle of one's own happiness, or by means of reason directed to objects of our possible volition in general, as in the principle of perfection, the will never determines itself *immediately* by the representation of the action. It is determined only by the incentive that the anticipated effect of the action exerts on the will: *I ought to do something, because I want something else*. And here yet another law must be laid down as the basis in my subject, according to which I necessarily will this other thing—and this law in turn needs an imperative to limit this maxim. For the impulse that the representation of an object possible through our powers is supposed to exert on the subject's will, in accordance with the subject's natural constitu-

tion, belongs to the nature of the subject—whether to its sensibility (inclination and taste) or to its understanding and reason, which by the particular constitution of their nature take satisfaction in exercising themselves on an object. So it would really be nature that gives the law. As such, this law not only must be cognized and proved through experience, and is therefore in itself contingent and consequently unfit to be an apodictic practical rule, as the moral rule must be; it is also *always only heteronomy* of the will. The will does not give itself the law; rather, an alien impulse gives it the law, by means of a nature of the subject attuned to its receptivity.

The absolutely good will, whose principle must be a categorical imperative, will therefore, being indeterminate with regard to all objects, contain merely the *form of willing* in general, and that as autonomy. That is, the fitness of the maxim of every good will to make itself a universal law is itself the sole law that the will of every rational being imposes on itself, without laying any incentive or interest at its basis.

How such a synthetic practical proposition is possible a priori, and why it is necessary, is a problem whose solution no longer lies within the bounds of the metaphysics of morals. Nor have we [445] asserted its truth here, much less claimed to have a proof of it in our power. We showed only, by developing the concept of morality that is generally in currency, that an autonomy of the will unavoidably attaches to it, or rather lies at its basis. Whoever holds morality to be something, and not a chimerical idea without truth, must therefore also admit the principle of it that we have put forward. This section, then, like the first, was merely analytic. That morality is no figment of the brain—which follows if the categorical imperative, and with it the autonomy of the will, is true and absolutely necessary as an *a priori* principle—requires a *possible synthetic use of pure practical reason*. But we cannot venture on this use without first giving a *critique* of this faculty of reason itself. In the last section we have to set out the main features of such a critique, as far as they suffice for our purpose.

Section Three

Transition from the metaphysics of morals to the critique of pure practical reason
Kant links morality to freedom and asks how a categorical imperative is possible.

THE CONCEPT OF FREEDOM IS THE KEY TO THE EXPLANATION OF THE AUTONOMY OF THE WILL

Kant argues that a free will and a will under moral laws are one and the same.

446 The *will* is a kind of causality belonging to living beings insofar as they are rational, and *freedom* would be that property of this causality by which it can be effective independently of alien causes *determining* it; just as *natural necessity* is the property of the causality of all non-rational beings by which they are determined to activity through the influence of alien causes.

This explanation of freedom is *negative*, and therefore unfruitful for gaining insight into its essence. But from it flows a *positive* concept of freedom, which is all the richer and more fruitful. The concept of causality carries with it the concept of *laws*, according to which, through something we call a cause, something else—namely the effect—must be posited. Hence freedom, although it is not a property of the will according to natural laws, is not for that reason lawless at all. Rather, it must be a causality according to immutable laws, though laws of a special kind; for otherwise a free will would be an absurdity. Natural necessity was a heteronomy of efficient causes, since every effect was possible only according to the law that something else determined the efficient cause to its causality. What else, then,

can freedom [447] of the will be but autonomy, that is, the will's property of being a law to itself? But the proposition "The will is in all its actions a law to itself" signifies only the principle of acting on no maxim other than one that can also have itself as a universal law for its object. This, however, is precisely the formula of the categorical imperative and the principle of morality. So a free will and a will under moral laws are one and the same.

If, therefore, freedom of the will is presupposed, morality together with its principle follows from it by mere analysis of its concept. The principle, however, is still always a synthetic proposition: an absolutely good will is one whose maxim can always contain itself, regarded as a universal law, within itself. For this property of the maxim cannot be found by analyzing the concept of an absolutely good will. Such synthetic propositions, however, are possible only in that both cognitions are bound together through their connection with a third thing in which both are to be found. The *positive* concept of freedom supplies this third thing, which cannot be, as it is with physical causes, the nature of the world of sense (in whose concept the concepts of something as cause, in relation to *something else* as effect, come together). What this third thing is, to which freedom points us and of which we have an idea *a priori*, cannot be indicated just yet; nor can the deduction of the concept of freedom from pure practical reason, and with it the possibility of a categorical imperative, be made comprehensible yet. That still needs some preparation.

FREEDOM MUST BE PRESUPPOSED AS A PROPERTY OF THE WILL OF ALL RATIONAL BEINGS

Kant argues that every rational being must act under the idea of its own freedom.

It is not enough that we ascribe freedom to our will, on whatever grounds, if we do not have sufficient grounds for attributing the very same freedom to all rational beings as well. For since morality serves as a law for us only as *rational beings*, it must also hold for all rational beings; and since it must be

derived solely from the property of freedom, freedom must also be proved to be a property of the will of all rational beings. Nor is it enough to demonstrate freedom from certain supposed experiences [448] of human nature (though this is in any case absolutely impossible, and it can be demonstrated only *a priori*). Rather, one must prove that it belongs to the activity of rational beings endowed with a will in general. Now I say: every being that cannot act otherwise than *under the idea of freedom* is, for that very reason, really free in a practical respect. That is, all laws that are inseparably bound up with freedom hold for it just as if its will had also been declared free in itself, and validly so in theoretical philosophy.[1] Now I maintain that we must necessarily lend to every rational being that has a will the idea of freedom as well, under which alone it acts. For in such a being we think of a reason that is practical, that is, that has causality with regard to its objects. Now one cannot possibly think of a reason that would consciously receive direction from elsewhere with regard to its judgments; for then the subject would attribute the determination of its power of judgment not to its reason but to an impulse. Reason must regard itself as the author of its principles, independently of alien influences. Consequently, as practical reason, or as the will of a rational being, it must be regarded by itself as free. That is, the will of such a being can be a will of its own only under the idea of freedom, and so, for practical purposes, such a will must be ascribed to all rational beings.

1 KANT'S NOTE. I take this route—assuming, as sufficient for our purpose, freedom only as laid down by rational beings merely *in idea* as the basis of their actions—so that I need not commit myself to proving freedom in its theoretical respect as well. For even if this latter question is left unsettled, the same laws that would bind a being that was really free still hold for a being that cannot act otherwise than under the idea of its own freedom. Here, then, we can free ourselves from the burden that weighs on the theory.

ON THE INTEREST ATTACHING TO THE IDEAS OF MORALITY

Kant confronts an apparent circle between freedom and morality and proposes a different standpoint.

We have finally traced the determinate concept of morality back to the idea of freedom. But we could not prove freedom to be something real, not even in ourselves and in human nature. We [449] saw only that we must presuppose it if we want to think of a being as rational and as endowed with consciousness of its causality with regard to actions, that is, with a will. And so we find that, on just the same ground, we must attribute to every being endowed with reason and will this property of determining itself to act under the idea of its freedom.

From the presupposition of these ideas, however, there also flowed the consciousness of a law of action: that the subjective principles of actions, that is, maxims, must always be adopted in such a way that they can also hold objectively, that is, universally as principles, and so can serve for our own universal legislation. But why, then, should I subject myself to this principle—and do so as a rational being in general, and thereby also subject to it all other beings endowed with reason? I will grant that no interest *impels* me to do so, for that would yield no categorical imperative. But I must still necessarily *take* an interest in it and see how this comes about. For this "ought" is really a "will" that holds for every rational being on condition that reason in it were practical without hindrance. For beings like us, who are also affected by sensibility, as incentives of another kind, and in whom what reason would do by itself alone does not always happen, that necessity of action is called only an "ought," and the subjective necessity is distinguished from the objective.

It seems, then, that in the idea of freedom we really only presupposed the moral law, namely the principle of the autonomy of the will itself, and could not prove its reality and objective necessity on its own. In that case we would indeed still have gained something quite considerable, in that we would at least

have determined the genuine principle more precisely than has usually been done. But with regard to its validity, and the practical necessity of subjecting oneself to it, we would have got no further. For suppose someone asked us why, then, the universal validity of our maxim as a law must be the limiting condition of our actions; on what we base the worth we assign to this way of acting—a worth supposed to be so great that there can be no higher interest anywhere; and how it comes about that a human being believes that through this alone he [450] feels his personal worth, compared with which the worth of an agreeable or disagreeable condition is to count for nothing. To such a questioner we could give no satisfactory answer.

We do indeed find that we can take an interest in a personal quality that carries with it no interest in our condition at all, provided only that it makes us capable of sharing in that condition should reason bring about its distribution. That is, we find that the mere worthiness to be happy can be of interest in itself, even without the motive of sharing in this happiness. But this judgment is in fact only the effect of the importance of moral laws, which we have already presupposed (when we separate ourselves from all empirical interest through the idea of freedom). But that we ought to separate ourselves from this interest, that is, regard ourselves as free in acting and yet hold ourselves subject to certain laws, in order to find merely in our own person a worth that could compensate us for the loss of everything that gives our condition its worth—and how this is possible, and hence *on what ground the moral law obligates*—this we cannot yet see in this way.

Here, it must be frankly admitted, a kind of circle shows itself, from which, it seems, there is no escape. We take ourselves to be free in the order of efficient causes in order to think of ourselves as under moral laws in the order of ends; and we afterwards think of ourselves as subject to these laws because we have ascribed freedom of will to ourselves. For freedom and the will's own legislation are both autonomy, and hence reciprocal concepts. But for that very reason one of them

cannot be used to explain the other or to supply its ground. At most it can serve, for logical purposes, to reduce apparently different representations of one and the same object to a single concept (as different fractions of equal value are reduced to their lowest terms).

One way out, however, still remains to us: namely, to inquire whether, when we think of ourselves through freedom as causes efficient *a priori*, we do not take up a different standpoint from when we represent ourselves in terms of our actions as effects that we see before our eyes.

TWO STANDPOINTS

Kant distinguishes the world of sense from the world of understanding, to which the will belongs.

Making this observation requires no subtle reflection; one may suppose that even the most common understanding can make it, though after its own fashion, through an obscure discrimination [451] of the power of judgment which it calls feeling. The observation is this: all representations that come to us without our choice (such as those of the senses) let us know objects only as they affect us, while what they may be in themselves remains unknown to us. Hence, as far as this kind of representation is concerned, even with the most strenuous attention and all the clarity that the understanding may ever add, we can attain only to cognition of *appearances*, never of *things in themselves*. Once this distinction has been made (perhaps merely by noticing the difference between representations given to us from elsewhere, in which we are passive, and those we produce solely from ourselves, in which we show our own activity), it follows of itself that behind the appearances we must admit and assume something else that is not appearance, namely things in themselves. Yet we resign ourselves to the fact that, since they can never become known to us except as they affect us, we cannot come any closer to them and can never know what they are in themselves. This must yield a distinction, however crude, between a

world of sense and the *world of understanding*. The former can vary a great deal according to differences of sensibility in the various observers of the world, while the latter, which underlies it, always remains the same. Even about himself—and indeed as far as the acquaintance a human being has with himself through inner sensation goes—he may not presume to cognize what he is like in himself. For since he does not, as it were, create himself, and acquires his concept of himself not *a priori* but empirically, it is natural that he can gather information even about himself only through inner sense, and consequently only through the appearance of his nature and the way his consciousness is affected. Yet beyond this constitution of his own subject, composed of nothing but appearances, he must necessarily assume something else lying at its basis, namely his I, as it may be constituted in itself. With regard to mere perception and receptivity to sensations, he must therefore count himself as belonging to the *world of sense*; but with regard to whatever in him may be pure activity (what reaches consciousness not through any affection of the senses but immediately), he must count himself as belonging to the *intellectual world*, of which, however, he knows nothing further.

The reflective person must draw a conclusion of this kind about all the things [452] that may present themselves to him. Presumably it is also to be found in the most common understanding. As is well known, the common understanding is much inclined to expect something invisible, active in itself, behind the objects of the senses; but it spoils this again by promptly making the invisible sensible once more, that is, by wanting to make it an object of intuition, and so it becomes not one degree wiser.

Now a human being really does find in himself a faculty by which he distinguishes himself from all other things, and even from himself insofar as he is affected by objects; and that faculty is *reason*. As pure self-activity, reason is raised even above the *understanding* in this respect. The understanding, too, is self-activity: it does not, like sense, contain merely representations

that arise only when we are affected by things (and so are passive). Yet from its activity it can produce no concepts other than those that serve merely *to bring sensible representations under rules* and thereby to unite them in one consciousness; without this use of sensibility it would think nothing at all. Reason, by contrast, shows under the name of ideas a spontaneity so pure that it goes far beyond anything sensibility can ever supply to it. It proves its highest function in distinguishing the world of sense and the world of understanding from each other, and thereby marking out the limits of the understanding itself.

For this reason a rational being must regard itself *as an intelligence* (hence not from the side of its lower powers) as belonging not to the world of sense but to the world of understanding. It therefore has two standpoints from which it can consider itself and recognize laws for the use of its powers, and consequently for all its actions: *first*, insofar as it belongs to the world of sense, under laws of nature (heteronomy); *second*, as belonging to the intelligible world, under laws which, being independent of nature, are not empirical but grounded merely in reason.

As a rational being, and hence one belonging to the intelligible world, a human being can never think of the causality of his own will otherwise than under the idea of freedom; for independence from the determining causes of the world of sense (which reason must always ascribe to itself) is freedom. Now the concept of *autonomy* is inseparably bound up with the idea of freedom, and with autonomy the universal principle of morality, which in idea lies at the basis of all actions [453] of *rational* beings, just as the law of nature lies at the basis of all appearances.

Now the suspicion we raised above has been removed: the suspicion that a hidden circle was contained in our inference from freedom to autonomy, and from autonomy to the moral law. We might, that is, have laid down the idea of freedom only for the sake of the moral law, in order afterwards to infer the law from freedom in turn; and so we could have given no ground at all for the law, but could only have put it forward as the begging

of a principle—one that well-disposed souls would gladly grant us, but that we could never establish as a demonstrable proposition.[2] For now we see that when we think of ourselves as free, we transfer ourselves into the world of understanding as members of it and recognize the autonomy of the will together with its consequence, morality; but when we think of ourselves as under obligation, we regard ourselves as belonging to the world of sense and yet at the same time to the world of understanding.

HOW IS A CATEGORICAL IMPERATIVE POSSIBLE?

Kant explains the moral ought through our membership in both the intelligible and the sensible world.

As an intelligence, a rational being counts itself as belonging to the world of understanding, and it calls its causality a *will* merely as an efficient cause belonging to that world. On the other side, however, it is also conscious of itself as a part of the world of sense, in which its actions are found as mere appearances of that causality. Yet the possibility of these actions cannot be understood from that causality, which we do not know; instead, those actions must be understood as determined by other appearances, namely desires and inclinations, and so as belonging to the world of sense. As a mere member of the world of understanding, then, all my actions would be in perfect conformity with the principle of the autonomy of the pure will; as a mere part of the world of sense, they would have to be taken as in complete conformity with the natural law of desires and inclinations, and hence with the heteronomy of nature. (The former would rest on the supreme principle of morality, the latter on that of happiness.) But *the world of understanding contains the ground of the world of sense, and hence also of its laws*; it is thus immediately legislating with regard to my will (which belongs wholly to the world of understanding), and must also be thought of as such. So, although on the other side I am like a being

2 "Begging of a principle" renders Kant's German for *petitio principii*, the logical fault of assuming what is to be proved.

belonging to the [454] world of sense, I shall nevertheless have to recognize myself, as an intelligence, as subject to the law of the former, that is, of reason, which contains that law in the idea of freedom, and so as subject to the autonomy of the will. Consequently I must regard the laws of the world of understanding as imperatives for me, and actions in conformity with this principle as duties.

And so categorical imperatives are possible because the idea of freedom makes me a member of an intelligible world. If I were only that, all my actions *would* always be in conformity with the autonomy of the will; but since I also intuit myself as a member of the world of sense, they *ought* to be in conformity with it. This *categorical* ought represents a synthetic proposition *a priori*, because to my will as affected by sensible desires there is added the idea of this same will, but as belonging to the world of understanding—a pure will, practical of itself, which contains the supreme condition of the former will according to reason. This is roughly how concepts of the understanding, which of themselves signify nothing but lawful form in general, are added to the intuitions of the world of sense and thereby make possible synthetic propositions *a priori*, on which all cognition of a nature rests.

The practical use of common human reason confirms the correctness of this deduction. There is no one—not even the worst villain, provided he is otherwise used to employing reason—who, when shown examples of honesty of purpose, of steadfastness in following good maxims, of sympathy and general benevolence (and these combined, moreover, with great sacrifices of advantage and comfort), does not wish that he too might be so disposed. He cannot quite bring this about in himself, but only because of his inclinations and impulses; and at the same time he wishes to be free of such inclinations, which are burdensome even to himself. He thereby proves that, with a will free from the impulses of sensibility, he transfers himself in thought into an order of things altogether different from that of his desires in the field of sensibility. For from that wish he can expect no grati-

fication of his desires, and hence no condition that would satisfy any of his actual or otherwise conceivable inclinations (for then the very idea that draws the wish from him would lose its preeminence); he can expect only a greater inner worth of his person. He believes himself to be this better person, [455] however, when he transfers himself to the standpoint of a member of the world of understanding—a standpoint to which the idea of freedom, that is, independence from *determining* causes of the world of sense, involuntarily constrains him. In it he is conscious of a good will which, by his own admission, constitutes the law for his evil will as a member of the world of sense—a law whose authority he recognizes even as he transgresses it. The moral ought is therefore his own necessary willing as a member of an intelligible world, and it is thought of by him as an ought only insofar as he at the same time regards himself as a member of the world of sense.

ON THE OUTERMOST BOUNDARY OF ALL PRACTICAL PHILOSOPHY

Kant argues that freedom cannot be explained, only defended, and marks the limits of practical philosophy.

All human beings think of themselves as free in their will. Hence all judgments about actions as ones that *ought to have been done*, even though they *were not done*. Yet this freedom is not a concept of experience, nor can it be one, because it always remains even though experience shows the opposite of those demands that are represented as necessary on the presupposition of freedom. On the other hand, it is just as necessary that everything that happens be inevitably determined according to laws of nature; and this natural necessity is not a concept of experience either, precisely because it carries with it the concept of necessity, and hence of a cognition *a priori*. But this concept of a nature is confirmed by experience, and must itself unavoidably be presupposed if experience—that is, cognition of the objects of the senses connected according to universal laws—is to be possible. Freedom is therefore only an *idea* of reason,

whose objective reality is in itself doubtful; nature, however, is a *concept of the understanding*, which proves its reality in examples from experience, and must necessarily do so.

Now a dialectic of reason arises from this, since with regard to the will the freedom ascribed to it seems to stand in contradiction with natural necessity. At this fork in the road reason, *for speculative purposes,* finds the path of natural necessity far better trodden and more serviceable than that of freedom. Yet *for practical purposes* the narrow footpath of freedom is the only one on which it is possible, in what we do and leave undone, to make [456] use of our reason. Hence it is just as impossible for the subtlest philosophy as for the most common human reason to argue freedom away. Philosophy must therefore presuppose that no true contradiction will be found between freedom and natural necessity in the very same human actions; for it can no more give up the concept of nature than the concept of freedom.

Nevertheless, this apparent contradiction must at least be eradicated in a convincing way, even if we could never comprehend how freedom is possible. For if the very thought of freedom contradicted itself, or contradicted nature, which is just as necessary, freedom would have to be given up altogether in favor of natural necessity.

But it would be impossible to escape this contradiction if the subject who fancies himself free were to think of himself *in the same sense, or in the very same relation,* when he calls himself free as when he takes himself to be subject to the law of nature with regard to the same action. It is therefore an inescapable task of speculative philosophy at least to show two things. First, that its illusion about the contradiction rests on our thinking of the human being in a different sense and relation when we call him free than when we hold him, as a part of nature, to be subject to its laws. Second, that the two not only *can* very well stand together, but must also be thought of *as necessarily united* in the same subject. For otherwise no ground could be given why we should burden reason with an idea which, although it can be united *without contradiction* with another, sufficiently estab-

lished idea, nevertheless entangles us in a business that drives reason into a very tight corner in its theoretical use. This duty, however, falls solely to speculative philosophy, so that it may clear a free path for practical philosophy. So it is not left to the philosopher's discretion whether he will remove the apparent conflict or leave it untouched. For in the latter case the theory on this point is *bonum vacans* [ownerless property], of which the fatalist can rightfully take possession, driving all moral philosophy out of what it supposed to be its property but held without title.[3]

Still, it cannot yet be said here that the boundary of practical philosophy begins. For the settlement of that dispute does not belong to practical philosophy at all; it only demands of speculative reason that it bring to an end the discord in which it entangles itself in theoretical questions, [457] so that practical reason may have peace and security from external attacks that could contest the ground on which it wants to build.

But the rightful claim even of common human reason to freedom of the will is grounded on the consciousness, and the admitted presupposition, of the independence of reason from merely subjectively determining causes, which together make up what belongs merely to sensation, and hence falls under the general name of sensibility. A human being who regards himself in this way as an intelligence thereby places himself in another order of things, and in a relation to determining grounds of an entirely different kind, when he thinks of himself as an intelligence endowed with a will, and consequently with causality, than when he perceives himself as a phenomenon in the world of sense (which he really is as well) and subjects his causality to external determination according to laws of nature. Now he soon becomes aware that both can take place at the same time—indeed, that they must. For there is not the slightest contradiction in a *thing in appearance* (belonging to the world of sense) being subject to certain laws of which the very same thing, as a

3 A term of Roman law for property without an owner, which falls to whoever first occupies it.

thing or being *in itself*, is independent. But that he must represent and think of himself in this twofold way rests, as regards the first, on the consciousness of himself as an object affected through the senses, and as regards the second, on the consciousness of himself as an intelligence, that is, as independent of sensible impressions in the use of reason (and hence as belonging to the world of understanding).

That is why a human being lays claim to a will that lets nothing be put down to its account that belongs merely to his desires and inclinations, and why, on the contrary, he thinks of actions as possible through it—indeed, as necessary—that can be done only by setting aside all desires and sensible stimuli. The causality of such actions lies in him as an intelligence, and in the laws of effects and actions according to principles of an intelligible world. Of that world he knows nothing more than that in it reason alone—and indeed pure reason, independent of sensibility—gives the law. Moreover, since there he is his proper self only as an intelligence (whereas as a human being he is only an appearance of himself), those laws apply to him immediately and categorically. So whatever inclinations and impulses (and hence the whole nature of the world of sense) incite him to cannot impair the laws of his willing as an intelligence [458] —so much so that he does not hold himself answerable for the inclinations and impulses and does not ascribe them to his proper self, that is, to his will, though he does ascribe to it the indulgence he might show toward them if he allowed them influence on his maxims to the detriment of the rational laws of the will.

Practical reason does not overstep its boundaries at all by *thinking* itself into a world of understanding; it would do so only if it wanted to *intuit* or *feel* itself into it. The former is only a negative thought with respect to the world of sense, which gives reason no laws in determining the will. It is positive only in this single point: that this freedom, as a negative determination, is at the same time bound up with a (positive) faculty, and indeed with a causality of reason that we call a will—a faculty of acting in such a way that the principle of one's actions conforms to the

essential character of a rational cause, that is, to the condition that the maxim be universally valid as a law. But if reason were also to fetch an *object of the will*, that is, a motive, from the world of understanding, it would overstep its boundaries and presume to know something of which it knows nothing. The concept of a world of understanding is therefore only a *standpoint* that reason sees itself compelled to take outside appearances *in order to think of itself as practical*. This would not be possible if the influences of sensibility were determining for the human being; but it is nonetheless necessary, unless he is to be denied consciousness of himself as an intelligence, and hence as a rational cause, active through reason, that is, operating freely. This thought certainly brings with it the idea of another order and legislation than that of the mechanism of nature, which applies to the world of sense; and it makes necessary the concept of an intelligible world (that is, the whole of rational beings as things in themselves). But it does so without the slightest presumption to think of that world here any further than merely according to its *formal* condition—that is, the universality of the will's maxim as law, and hence the will's autonomy, which alone is consistent with its freedom. All laws that are directed to an object, by contrast, yield heteronomy, which is found only in laws of nature and can apply only to the world of sense.

But reason would then overstep all its bounds if it presumed to *explain* how pure reason can be [459] practical—which would be exactly the same task as explaining *how freedom is possible*.

For we can explain nothing except what we can trace back to laws whose object can be given in some possible experience. Freedom, however, is a mere idea, whose objective reality can in no way be shown according to laws of nature, and hence not in any possible experience either. Since no example of it can ever be supplied by any analogy, it can never be comprehended, or even so much as understood. It holds only as a necessary presupposition of reason in a being that believes itself conscious of a will, that is, of a faculty distinct from the mere faculty of desire (namely, the faculty of determining itself to act as an

intelligence, and hence according to laws of reason, independently of natural instincts). But where determination according to laws of nature ceases, all *explanation* ceases too, and nothing remains but *defense*: that is, repelling the objections of those who claim to have seen deeper into the essence of things and therefore boldly declare freedom to be impossible. One can only show them that the contradiction they suppose they have discovered here lies nowhere else than in this. In order to make the law of nature hold with regard to human actions, they had to regard the human being as an appearance; and now, when they are asked to think of him, as an intelligence, also as a thing in itself, they still go on regarding him as an appearance even there. In that case, admittedly, separating his causality (that is, his will) from all the natural laws of the world of sense in one and the same subject would be a contradiction. But the contradiction falls away if they would only reflect and admit, as is fair, that behind appearances there must after all lie, as their ground, the things in themselves (though hidden), and that one cannot demand that the laws by which these things operate be the same as those to which their appearances are subject.

The subjective impossibility of *explaining* the freedom of the will is the same as the impossibility of discovering an *interest*[4] that a human being could take in moral laws, and of making it 460 comprehensible. And yet he really does take an interest in them. The basis of this interest in us we call moral feeling. Some have wrongly passed this feeling off as the standard of our moral

4 KANT'S NOTE. Interest is that by which reason becomes practical, that is, becomes a cause determining the will. That is why we say only of a rational being that it takes an interest in something; creatures without reason feel only sensible impulses. Reason takes an immediate interest in an action only when the universal validity of its maxim is a sufficient determining ground of the will. Only such an interest is pure. But if reason can determine the will only by means of another object of desire, or on the presupposition of a particular feeling of the subject, then reason takes only a mediate interest in the action. And since reason by itself, without experience, can discover neither objects of the will nor a particular feeling underlying the will, this latter interest would be only empirical, and not a pure interest of reason. The logical interest of reason (in advancing its insights) is never immediate, but presupposes purposes for which reason is to be used.

judgment, whereas it must rather be regarded as the *subjective* effect that the law exerts on the will, while reason alone supplies the objective grounds for it.

In order to will that for which reason alone prescribes the ought to a sensibly affected rational being, there must of course be a capacity of reason to *instill a feeling of pleasure* or of delight in the fulfillment of duty, and hence a causality of reason to determine sensibility in accordance with its principles. But it is altogether impossible to see—that is, to make comprehensible *a priori*—how a mere thought, which itself contains nothing sensible, can produce a sensation of pleasure or displeasure. For that is a particular kind of causality, about which, as about all causality, we can determine nothing at all *a priori*; for this we must consult experience alone. But experience can supply no relation of cause to effect except as between two objects of experience; whereas here pure reason, through mere ideas (which yield no object at all for experience), is supposed to be the cause of an effect that does indeed lie in experience. Hence it is quite impossible for us human beings to explain how and why the *universality of the maxim as law*, and hence morality, interests us. This much only is certain: it is not valid for us *because it interests us* (for that is heteronomy, and a dependence of practical reason on sensibility, namely [461] on an underlying feeling, in which case it could never be morally legislating). Rather, it interests us because it is valid for us as human beings, since it has sprung from our will as an intelligence, and hence from our proper self; *but what belongs to mere appearance is necessarily subordinated by reason to the character of the thing in itself*.

The question, then, of how a categorical imperative is possible can indeed be answered to the extent that one can state the sole presupposition under which alone it is possible, namely the idea of freedom, and also to the extent that one can see the necessity of this presupposition. That suffices for the *practical use* of reason, that is, for conviction of the *validity of this imperative*, and hence also of the moral law. But how this presupposition itself is possible can never be seen by any human reason.

On the presupposition of the freedom of the will of an intelligence, however, its *autonomy*, as the formal condition under which alone the will can be determined, is a necessary consequence. To presuppose this freedom of the will is not only quite *possible* (as speculative philosophy can show), without falling into contradiction with the principle of natural necessity in the connection of appearances in the world of sense. It is also, without any further condition, *necessary* for a rational being conscious of its causality through reason, and hence of a will (which is distinct from desires), to lay freedom practically—that is, in idea—at the basis of all its voluntary actions as their condition. But *how* pure reason can be practical of itself, without other incentives that might be taken from somewhere else—that is, *how* the mere *principle of the universal validity of all its maxims as laws* (which would certainly be the form of a pure practical reason), without any matter (object) of the will in which one might take some interest in advance, can of itself furnish an incentive and produce an interest that would be called purely *moral*; or, in other words, *how pure reason can be practical*—to explain this, all human reason is wholly incapable, and all the effort and labor spent seeking an explanation of it is lost.

It is just the same as if I sought to fathom how freedom itself is possible as the causality of a will. For there I abandon the philosophical [462] ground of explanation, and I have no other. I could, of course, now roam about in the intelligible world that still remains to me, the world of intelligences. But although I have an *idea* of it, and one with good grounds, I have not the slightest *knowledge* of it, nor can I ever attain such knowledge through all the striving of my natural faculty of reason. It signifies only a something that remains when I have excluded from the determining grounds of my will everything that belongs to the world of sense—merely in order to restrict the principle of motives drawn from the field of sensibility, by setting bounds to it and showing that it does not comprise all in all within itself, but that there is still more beyond it. This more, however, I know nothing further about. Of the pure reason that thinks this

ideal, once all matter—that is, cognition of objects—has been set aside, nothing remains to me but the form: namely, the practical law of the universal validity of maxims, and, in conformity with it, the thought of reason, in relation to a pure world of understanding, as a possible efficient cause, that is, as a cause determining the will. The incentive must be entirely lacking here—unless this idea of an intelligible world were itself the incentive, or that in which reason originally took an interest. But to make this comprehensible is precisely the problem we cannot solve.

Here, then, is the uppermost boundary of all moral inquiry. Determining it is of great importance for this reason too: so that reason, on the one hand, does not search about in the world of sense, in a way harmful to morals, for the supreme motive and for an interest that is comprehensible but empirical; and so that, on the other hand, it does not flap its wings powerlessly, without moving from the spot, in the space of transcendent concepts—empty for it—that goes by the name of the intelligible world, and lose itself among chimeras. For all that, the idea of a pure world of understanding, as a whole of all intelligences to which we ourselves belong as rational beings (although on the other side we are at the same time members of the world of sense), always remains a usable and permitted idea for the purposes of a rational faith, even though all knowledge ends at its boundary. Through the glorious ideal of a universal kingdom of *ends in themselves* (rational beings)—to which we can belong as members only when we carefully [463] conduct ourselves according to maxims of freedom as if they were laws of nature—it serves to produce in us a lively interest in the moral law.

CONCLUDING REMARK

Kant concludes that we cannot comprehend the moral imperative's necessity, only its incomprehensibility.

The speculative use of reason *with regard to nature* leads to the absolute necessity of some supreme cause *of the world*; the practical use of reason *with a view to freedom* also leads to absolute

necessity, but only *of the laws of the actions* of a rational being as such. Now it is an essential *principle* of all use of our reason to push its cognition to the point of consciousness of its *necessity* (for without this it would not be cognition by reason). But it is an equally essential *limitation* of this very same reason that it can see the *necessity* neither of what exists or happens, nor of what ought to happen, unless a *condition* under which it exists, or happens, or ought to happen, is laid at its basis. In this way, however, the constant inquiry after the condition only postpones reason's satisfaction further and further. Hence reason restlessly seeks the unconditionally necessary and sees itself compelled to assume it without any means of making it comprehensible to itself—happy enough if it can only find the concept that is compatible with this presupposition. It is therefore no objection to our deduction of the supreme principle of morality, but a reproach that would have to be made against human reason in general, that it cannot make an unconditional practical law (such as the categorical imperative must be) comprehensible as regards its absolute necessity. For reason cannot be blamed for being unwilling to do this through a condition, namely by means of some interest laid at its basis, since then the law would not be moral, that is, the supreme law of freedom. And so, while we do not indeed comprehend the practical unconditional necessity of the moral imperative, we do comprehend its *incomprehensibility*. And that is all that can fairly be asked of a philosophy that strives in its principles to reach the very limit of human reason.