

Marcus Willaschek

## Practical Reason

A commentary on Kant's *Groundwork of the Metaphysics of Morals* (GMS, 412–417)<sup>1</sup>

### 1. Introduction

On pages 412–417 of the *GMS*, Kant introduces his conception of the will as *practical reason* and the closely related distinction between hypothetical and categorical imperatives in a rather concise manner. In the following, I will address the much discussed definition of the will (4: 412), the concepts of both the holy will (4: 412 und 414) and the imperative (4: 413–414), as well as the differences among the different kinds of imperatives (4: 414–417), while explaining Kant's central concepts and theses. Prior to this, however, I will present a short overview of Kant's general line of argument.

### 2. Overview

Practical reason is the ability to act rationally – that is, the ability of a person to rationally coordinate her goals and ends and to orient her actions according to these rationally set ends. Kant identifies this ability with the *will*, or with the ability of a rational being (more precisely, a being possessing reason) to be the cause of its conduct through its own “representations” (mental states). In other words, having a will means being able to act in accordance with one's own rational representations.

Within the concept of practical reason, Kant distinguishes *pure* from *empirically qualified* practical reason – a distinction that he first

<sup>1</sup> For their valuable advice and criticism, I would like to thank Alexander Bagattini, Steffi Schadow, Dieter Schönecker as well as the participants of the preparatory conference in Bonn.

makes explicit in the *KpV*, and that is already anticipated (and implicitly presupposed) in the *GMS* (cf. 4: 389, where Kant speaks of a "practical pure reason" as well as of the question of how "pure reason can be practical," 4: 458; cf. 461). God and the angels (if they exist) possess exclusively *pure* practical reason; their decisions are in no way influenced by subjective factors, and particularly not by sensible (*sinnliche*) inclinations, but solely by objective, and specifically moral, rational grounds. Each of their wills is a "holy will," which *necessarily* and *exclusively* seeks the moral good.

Humans, on the other hand (as the only rational beings of whose existence we have knowledge), are influenced by their naturally and socially conditioned wishes and inclinations in many ways. Their wills are, therefore, not holy, for when it accommodates their interests they can decide to act immorally. Their practical reason is therefore *empirically qualified*: as instrumental rationality, it serves to satisfy empirically-given inclinations in the most effective way possible. If, on the other hand, as Kant supposes, moral principles hold without exception and necessarily for all beings who possess reason, thus also for humans, then it must be rational for every rational being to follow these principles, independent of its respective inclinations or wishes. Accordingly, humans must not only have empirically qualified reason, but also *pure* practical reason; that is, they must be able to orient their action according to moral principles when this would not further, or would even contradict, their own interests.

The human will is thus distinguished from a holy will in two ways: first, our decisions do not necessarily accord with that which would be rational to do, and second, "rational" for us does not only mean "morally good," but also "useful for the satisfaction of subjective inclinations." The first difference leads to the concept of an imperative, the second to the distinction between hypothetical and categorical imperatives.

Imperatives are propositions in which what is good and reasonable to do is expressed in the form of a command. Such commands are not directed at a being with a holy will that necessarily does what is rational, but rather at beings such as humans, who *can* act rationally but who do not necessarily do so (i. e., those for whom both rational and irrational actions are possible in any given situation). Most people are thus capable of doing something because they have recognized it as rational (e. g. taking care of their health or paying their taxes), but this rational insight does not necessarily motivate them to a corresponding course of action, because their inclinations (e. g. towards indolence) or their self-interest may oppose it. The principles of rational action, therefore,

appear for human beings as *commands* that express themselves in an *ought*: "(From a rational point of view) you ought to live healthily" or "*ought*:" "(From a rational point of view) you ought not cheat on your taxes." This "ought" is not an expression on the part of a power or convention (as is the case when one says that someone *ought* to follow the command of a superior or *ought* to greet one's neighbors), but an expression of a specifically *rational* form of motivation, which Kant describes as rational "necessitation" (*Nötigung*). This necessitation serves as a motive for doing what is rational even when we find it difficult (and must therefore "force" ourselves to do it). The more rational one is in a practical respect, the more fully developed will this motive be.

By prescribing an action, imperatives distinguish it as in some way "good." Corresponding to the distinction between a pure and a practically qualified reason, there are also two kinds of imperatives: those that demand an action because it is "good" as a *means* to the effective satisfaction of inclinations and needs, and those that prescribe an action because it is "rational per se" or "good in itself," without taking into account the ends of the agent. The former are hypothetical imperatives and the latter, categorical imperatives. Kant divides the hypothetical imperatives again into two classes, depending on whether they prescribe an action as a means to a *possible* or an *actual* end. Thus Kant arrives at a tripartite division of imperatives, giving several alternative descriptions of each of the three classes, depending on which of their characteristics he wants to emphasize: (1) problematically practical principles (= imperatives/rules of skill, technical imperatives), (2) assertorically practical principles (= rules/counsels of prudence, pragmatic imperatives), (3) apodictically practical principles (= imperatives/laws of morality, moral imperatives). (1) and (2) are the hypothetical, (3) the categorical imperatives.

### 3. *The definition of the will (4: 412)*<sup>2</sup>

Kant defines the will as the "capacity to act in accordance with the representation of laws, that is, in accordance with principles," and he identifies this capacity with practical reason. This passage has given rise to a variety of different interpretations (cf. Laberge 1989; Tim-

<sup>2</sup> The following translations of passages from Kant's works are based on the Cambridge Edition of the Works of Immanuel Kant (New York: Cambridge University Press, 1992-).

mermann 2003, § 9), and even if one takes the context into account the passage remains ambiguous. Kant begins with this observation:

- (1) Everything in nature works in accordance with laws.

This claim follows from the regularity conception of causation that Kant inherited from Hume (cf. Willaschek 1992, 34–43). According to this view,  $a$  at time  $t_0$  is the cause of  $b$  at time  $t_1$ , if  $a$  is of type  $A$  and  $b$  is of type  $B$ , such that from things or events of type  $A$  at  $t_0$  there always follow things or events of type  $B$  at  $t_1$ . In that case, “If  $A$  at  $t_0$ , then  $B$  at  $t_1$ ” is a universally valid law. Thus the claim holds analytically that all things in nature “work in accordance with laws,” or that their causality functions according to laws of nature.

- (2) Only a rational being has the **capacity** to act in accordance with the representation of laws, that is, in accordance with principles, or has a will.

The will is thus a **capacity** (i. e., an ability) of a rational being to act in a specific way, namely “in accordance with the representation of laws.” But what does it mean to act “in accordance with the representation of laws”? Which laws are meant here? It is exactly with regard to this question that interpretations diverge.

In order to understand Kant correctly here, it will be helpful to review the connections among the concepts *life*, *faculty of desire* (*Begehrungsvermögen*), and *will* (cf. with regard to this and what follows Willaschek 1992, 82–90). “*Life*,” as Kant writes in the *KpV*, “is the faculty of a being to act in accordance with the laws of the faculty of desire. The *faculty of desire* is the faculty of this being to be, through its representations, the cause of the actuality of the objects of these representations” (5: 9; cf. already in 2: 327 as well as 4: 544). As is briefly stated in the *MdS*, “The capacity of a being to act in accordance with its representations is called *life*” (6: 211). Here, “to act” has the broad meaning of the Latin “*agere*”: it means the same as “to bring about” or “to cause” (cf. Gerhardt 1986). Kant understands life as the purposeful spontaneous activity (*Selbsttätigkeit*) of an organism (or in the case of a “living God,” cf. 3:421, the activity of a non-material entity) in accordance with the Aristotelian-Leibnizian concept of *entelechia*. This spontaneous activity, however, is conceived (at least in the case of physical beings) in a distinctively “modern” way as a lawful and purely causal association. A living being also “works” according to laws: it has a representation of what it desires and this representation motivates it to act in a way that leads to the realization of the

desired object or goal. The causal connection between representation and bodily movement has the form of a natural law; more specifically, it conforms to a psychophysical lawfulness, a “law of the faculty of desire.” Because the desire comes from the living being itself, Kant can also say that it acts “from an inner principle” (4: 544).

As living beings, rational beings act or work “according to laws” – and specifically according to those of the (rational) faculty of desire. At the same time, they act “according to their representations,” and thus “according to the representation of laws.” But the laws according to the representation of which rational beings act, according to the *GMS*, are not the “laws of the faculty of desire” according to which they “act” or “work” as living beings. As we have seen, the latter are causal laws that describe the regular connection between the mental states (“representations”) and bodily movements of a living being. Like all natural laws they hold independently of whether or not a particular living being represents them to itself. Even if Kant’s formulation might seem to suggest something else, the laws mentioned in sentence (1) are of a completely different kind than the laws of which Kant speaks in sentence (2).

But what then are these laws “according to the representation of which” rational beings act? The answer does not unambiguously emerge from the aforementioned passage, but the context lends credence to the supposition that they are those “objective laws” of reason which are discussed at the end of the paragraph and in the paragraphs that follow (4: 413). What is of concern here is not only the moral law and other, more specific laws of morality, but also the “laws” of instrumental action, for, as Kant says, “All imperatives... indicate... the relation of an objective law of reason to a will that because of its subjective constitution is not necessarily determined by it” (4: 413, my emphasis). Even though Kant usually understands “practical laws” to be moral laws (cf. 5: 19), it seems that a broader reading is appropriate in this context: the laws according to whose representation only rational beings can act are the laws of rational action, whether they be laws of skill, pragmatic laws, or moral laws.

Kant calls the laws themselves (but not their “representation”) “principles” or “objective principles” (cf. e. g. 4: 400 note, 4: 413). The addition “that is, according to principles” can consequently not refer to the entire preceding expression, “according to the representation of laws,” for laws or principles are something quite different from representations of laws or principles. If the explanation beginning with “that is” (“d. i.”) refers only, however, to the expression “laws,” then

the repetition of the words "according to" ("nach") creates difficulties, for it would then be both the representation of laws and the laws (principles) themselves according to which rational beings act. This difficulty is resolved, however, if one considers what it means for Kant that a rational living being acts "according to a principle": it has a representation of the principle in question, which motivates it to realize the represented object (the principle). Someone who acts according to the representation of a law (a principle), *ipso facto* acts according to the law (principle) itself. Sentence (2) should then be understood in the following way:

- (2) Only rational beings can orient their behavior according to the representation of the laws of rational action, and thus act according to these laws or principles themselves.

One can also say this more concisely: only rational beings have the ability to act *rationally*, in both the instrumental and moral sense. Kant calls this ability for *rational* action and effort the "will." The human will, as one can also say with Kant, is a rational "faculty of desire in accordance with concepts" (6: 213); the "representations" through which we become cause of the actuality of represented objects are not (only) of an intuitive and sensible kind, but are (also) conceptual. Kant calls a concept that is at the same time the cause of the object represented through the concept an "end" (5: 180).

While the definition of the "will" as the rational faculty of desire corresponds with the philosophical tradition, and especially with the vocabulary of the Leibnizian-Wolffian school of philosophy (cf. e.g. Baumgarten, *Metaphysica* § 690), in the sentence that follows Kant chooses an at least terminological path of his own:

- (3) Since *reason* is required for the derivation of actions from laws, the will is nothing other than practical reason.

The expression "practical reason" takes up the Aristotelian language of *nous praktikos*, but has no modern-language predecessor in the 18th century (cf. Beck 1974, 265, note 14). Hence Kant was free to define this expression according to his own needs. Nevertheless, Kant apparently does not want to give a stipulative definition, as he justifies the identification of the will and practical reason with the remark that "*reason* is required for the derivation of actions from laws." One might grant Kant that "derivation" is the task of reason; at another point he expressly identifies reason with the "faculty of the determination of the particular through the universal (the derivation of prin-

ciples)" (20: 201). However, that means only that one cannot have a will without also having reason (cf. 4: 427: "The will is thought as a capacity to determine itself to action in conformity with the *representations of certain laws*. And such a capacity can be found only in rational beings.") This does not mean, however, that the will "is nothing other" than practical reason. For this, rather, two claims would have to hold, claims that Kant implicitly presupposes: first, that the will (as the capacity to act according to the representation of laws of reason) "is nothing other" than the capacity of the "derivation of actions from laws," and second, that for this capacity, practical reason is not only necessary (Kant: "is required"), but also sufficient.

The talk of a "derivation of actions" refers to Aristotle's practical syllogism in which a general major premise ("Everything sweet must be tasted") and a minor premise ("This is sweet") are followed directly by the execution of the rationally commanded action (cf. EN 1147a 25 ff.). Kant also assumes that practical reason does not only consist in deriving *propositions* or *opinions* about actions but also the *actions* themselves: *practical* reason is the ability to orient one's behavior according to the laws of reason. That is, it is not only the ability to *recognize* what is rational (the so-called "principlium diiudicationis"), but to also *do* it ("principlium executionis"). However, Kant's conception differs from the Aristotelian view in one important respect: practical reason does not produce the "derived" actions directly (i.e., simply through insight into their rationality), but indirectly, via a feeling. This feeling, however, which in the case of action according to moral laws Kant terms "respect for law" (cf. 4: 400, 440) or "moral feeling" (4: 460), is itself of rational origin, for it arises through insight into the rationality of moral laws (cf. 4: 460, 5: 76). In Kant's view practical reason delivers both objective reasons and the subjective "incentives" of rational action and is thus nothing other than the will or the ability to act rationally.

This identification of practical reason and the will raises a problem, however, when Kant directly afterwards (and in many other passages) speaks of reason "determining" the will (more or less effectively). How can this be if the two are identical? The answer lies in the ambiguity of the expressions "will" and "practical reason" (cf. on this and the following Willaschek 1992, 48–53). In the *GMS*, as we have already seen, Kant defines "will" and "practical reason" as the ability to act rationally. What is at issue there, however, is a complex ability. It includes on the one hand the ability to bring the totality of one's wishes and convictions into a rational order (which guarantees

consistency and completeness). Kant calls this ability *reason* and, insofar as it concerns primarily wishes, ends, and grounds for action, *practical reason* or *will*: "The will is therefore the faculty of desire considered not so much in relation to action (as choice (*Willkür*) is), but rather in relation to the *determining ground* of this choice of action. The will itself actually has no determining ground, but insofar as it can determine choice it is instead practical reason itself" (6: 213, my emphasis). In this passage from *MdS* Kant does not understand "will" and "practical reason" as he does in *GMS* in the broader sense, as the ability to act rationally, but rather in the narrower sense, as the ability to produce rational determining grounds for action. On the other hand, practical reason in the broad sense also includes the ability to actually *act* according to these rational grounds of action. This ability is what Kant calls "free choice" (*freie Willkür*). Choice is the capacity to act according to one's own representations, which are connected with pleasure and displeasure (*Lust* and *Unlust*) (6: 413); choice is *free* insofar as pleasure and displeasure do not determine behavior but rather serve as a basis for the rational establishing of ends (cf. *KrV* A 802/B 830 and, with a slight change of emphasis, 6: 413).

Practical reason or will in the *broader sense* as the ability to act rationally includes, first, practical reason or the will in the *narrower sense* (that is, the ability to bring one's wishes and ends into a consistent scheme of grounds for action) and, second, "free choice" (that is, the ability to *do* what is rationally desired). The will, choice, and practical reason are therefore not distinct causal instances or subject-like homunculi, but instead aspects of the complex ability to act rationally. When Kant speaks of the fact that reason *determines* the will or choice, he means that practical reason in the narrower sense (the capacity to set rational ends) provides the "determining grounds" for the will in the broader sense (the ability to act *rationally*) or free choice (the ability to *act* rationally). That reason determines the will thus means – depending on emphasis – either that one actually does what is rationally desired, or that what one does is actually rational. When Kant, however, equates reason and the will, then he means either (in the broader sense) the ability to act rationally or (in the narrower sense) the ability to establish ends rationally. Although Kant's use of words is confusing and not always uniform, his conception of the will as practical reason is in fact uniform and thoroughly comprehensible.

A further meaning of the term "will" in Kant has not yet been mentioned: namely, according to Kant the will is not only the ability

to act according to laws or principles and thus rationally, but rather, as *free will*, also the *law-giving* faculty (cf. 6: 226). This means that a free rational being can not only orient its decisions and actions according to the rational principles that it is given and adherence to which serves the satisfaction of subjective inclinations, but that it also has the ability to "give" itself these principles according to which it acts and decides. Kant calls this ability "autonomy" or self-legislation (*Selbstgesetzgebung*) (4: 440ff.). The fact that a rational being "gives" itself the principles according to which it acts is obviously not intended to mean that the recognition of these principles is the result of a conscious and deliberate decision that could also have turned out otherwise. It means that the rationality of a rational agent consists in the recognition of these principles; they are not given to her by a foreign authority (heteronomy), but are rather an expression of her own reason. While Kant in the *GMS* takes a heteronomous will to be conceivable (4: 441), in the *MdS* he seems to almost identify the will with the ability of self-legislation (6: 226).

#### 4. The concept of a holy will (4: 412 and 414)

Kant's transition from the definition of the will to his conception of the imperative is constituted by the concept of the *holy will*, which he introduces in two steps: "If reason infallibly determines the will, the actions of such a being that are recognized as objectively necessary are also subjectively necessary, that is, the will is a capacity to choose only that which reason independently of inclination recognizes as practically necessary, that is, as good." (4: 412). Soon afterwards, this necessarily good will turns out to be a "holy will": "A perfectly good will would, therefore, stand just as much under objective laws (of the good), but it could not on this account be represented as *necessitated* to actions in conformity with law, since it of itself, according to its subjective constitution, can only be determined by the representation of the good. Hence no imperatives hold for the divine will and in general for a holy will: the 'ought' is out of place here, because volition is already of itself necessarily in accord with the law" (4: 414). A holy will is thus a will which "of itself, according to its subjective constitution" can do nothing other than the rational and good. This conception raises a number of questions, only two of which I want to consider more closely here: (1) Is the conception of a will that wills "only" the objectively rational, i.e., the morally good, consistent? (2) And how

does this concept of a “perfectly good will” relate to that of the “good will” that is the topic of the beginning of the first section?

(1) A holy will, in contrast to a human will, is a “capacity to choose *only* that which reason independently of inclination” recognizes as good; such a being “can be determined *only* through the representation of the good.” These formulations suggest that a being can only have a holy will if its will is not “affected” by inclinations, but instead acts *exclusively* according to the objective principles of reason. This becomes clear, too, through comparison with the human will for which, in contrast to the holy will, reason alone “does not adequately determine the will” and thus which “is also subject to subjective conditions (certain incentives)” (4: 412). Thus the holy will would be subject to *no* subjective conditions, such that reason *alone* would determine it sufficiently. Similarly, Kant writes in the *KpV* that the moral law has for humans “the character of an imperative, because in them, as rational beings, one can presuppose a *pure* will but, as beings affected by needs and sensible motives, not a holy will, that is, a will that would be incapable of maxims conflicting with the moral law” (5: 32). In this passage as well it sounds as if a being with a holy will may not be affected by “needs and sensible motives,” because it is this state of being affected that prevents humans from attaining this holy will.

Such a view would be awkward, however, since a being without “inclinations” would be incapable of action. Rational principles alone are not sufficient to establish how one should act in a certain situation. With regard to the principle of instrumental action (“He who wills the end, also rationally wills the necessary means to that end”) this is obvious: it presupposes that one already has ends; what these ends are does not only depend on rational considerations, but also on that to which one has an “inclination.” Even moral principles, and above all the categorical imperative, are by themselves not sufficient to be taken to show a certain action to be rationally commanded. After all, one is supposed to test, using the categorical imperative, if the maxim by which one intends to act could be a universal law. Maxims, however, are *subjective* rules of action, which “reason determines according to the conditions of the *subject* (more often his ignorance or also his inclinations)” (4: 421; my emphasis). A being that has no inclinations (and is also not subject to any other “subjective conditions”) has also, therefore, no maxims, and thus nothing to which it could apply the categorical imperative. Thus, even a being with a

holy will needs inclinations (or other not purely rational “subjective conditions”) in order to be able to act rationally.

Now in the passage cited above, Kant does not explicitly exclude the possibility that a being with a holy will could have inclinations, but instead says only that inclinations may have no influence on its will. As we have just seen, however, even this is problematic: a “capacity to choose *only* that which reason independently of inclination recognizes as practically necessary, that is, as good,” is not enough to decide on concrete actions and would therefore not be a will. Kant’s characterization of the holy will as completely independent of inclinations is thus too strong. However, what is important for Kant in the concept of the holy will is not its independence from inclinations, but its necessary correspondence with the laws of reason and of the good: “The will whose maxims necessarily harmonize with the laws of autonomy is a *holy*, absolutely good will” (4: 439). And for this necessary correspondence, in contrast to what Kant’s formulations suggest, a complete independence from inclinations is not necessary. A being whose will is capable of being affected by inclinations, but whose inclinations *necessarily* (because of its intrinsic constitution) correspond to the laws of reason, would have maxims, but “would not be capable of any maxims conflicting with the moral law” (5: 32) and would thus have a holy will.

In fact, many passages speak for the thesis that Kant was of the opinion that a holy will could well be affected by inclinations as long as these necessarily corresponded with the laws of reason. Kant writes in the *MdS* that in the case of a “holy (superhuman) being... no hindering impulses would impede the law of its will” (6: 405). And in lecture notes it is stated, “holiness is the absolute or unlimited moral perfection of the will. A holy being must not be affected by the least inclination *against morality*” (29: 1075). The addition “against morality” obviously only makes sense if it is at least conceivable that a holy being is at all affected by inclinations. Kant continues: “It must be impossible for him to want anything that would be contrary to the moral laws. Understood thusly, no being aside from God is holy; as every creature always has some needs, and if it wants to satisfy them, then inclinations as well, which *do not in each case correspond with morality*” (29: 1075). The holiness of a will thus consists in its *necessary* correspondence with the laws of reason and morality. This does not exclude the possibility that a being with a holy will is affected by inclinations, but only the possibility that it is affected by inclinations that could motivate it to do something

irrational and immoral. (The inclinations of a divine being – divine love, for example – would not be sensibly conditioned like ours, but would have non-sensible origins.)

(2) This takes us to the second question: How does the concept of the holy will relate to that of the good will, with which Kant's *Grundlegung* begins? Obviously, a holy will is also a good will, indeed even a "completely good" will (4: 414). But is the good will also always a holy will? Initially it seems plausible to assume that a good will does not necessarily have to be a holy will, but that a human will can also be good. However, upon looking more closely it is anything but clear how a good will differs from a holy will and thus whether humans can ever possess a good will.

Kant characterizes the good will as the only thing that is good "without limitation" (4: 393) or "in itself" (4: 394). It has an "absolute worth" (4: 394) that is independent of whether or not what is willed is also achieved.<sup>3</sup> What then constitutes a good will? Wood distinguishes between two possibilities: (1) the correspondence of the maxims with the moral law and (2) action from duty. In the first case the will of a being would be a good will if this being acted "in accordance with duty"; in the second case it would be good if it acted from duty. Because Wood understands action from duty as action from moral self-constraint, he prefers the first possibility. In fact it is implausible to assume that only a being which has to force itself to perform good deeds has a good will, for then God could not have a good will. But viewing the good will, with Wood, as a will whose maxims merely correspond to the moral law, and who thus acts *in accordance with duty*, one aspect of the good will to which Kant attached great importance is lost – namely, the *reliability* and *consistency* of the good will. A will is not good when it accidentally chooses the right maxims (4: 426) and thus acts in accordance with duty, but instead when it does this from an inner principle that guarantees action in accordance with duty under all circumstances. This principle is the moral law: as Kant says, for a being with a good will the "objective principle" (the moral law) serves "subjectively" as its "practical principle" (cf. 4: 400 note). In other words, a being with a good will acts according to the maxim to act only in accordance with such maxims that could be universal laws (cf. Willaschek 1992, 67ff.). Whoever acts according to this (meta)-maxim, acts out of "respect for the law" and thus "from duty," for

<sup>3</sup> Cf. on this subject Allen Wood's contribution to this volume.

"duty is the necessity of an action done out of respect for the law" (4: 400). The individual actions do not then accidentally correspond with the moral law; rather, they do so *necessarily*.

When Kant returns to the good will at the end of the second part of the *GMS*, it is exactly this necessity of correspondence with the moral law that he emphasizes: "We can now end where we set out from at the beginning – namely, with the concept of an unconditionally good will. That will is absolutely good *which cannot be evil*, hence whose maxim, if made into a universal law, *can never conflict with itself*. This principle is, accordingly, also its supreme law: act always on that maxim whose universality as a law you can at the same time will" (4: 437; my emphasis). A will is thus good if it *cannot* be evil, and therefore cannot be evil because it has made correspondence with the moral law into the "supreme law" or, as Kant says in another passage, into the "supreme condition of all maxims" (5: 31; cf. 6: 36).

If this interpretation of the concept of the good will is legitimate, then the question of the distinction between the good and the holy will becomes all the more urgent, since for the holy will Kant also emphasizes the necessity of correspondence with the moral law, and he equates the holy will with the "completely good will" (4: 414) and the "absolutely good will" (4: 439). Nevertheless, there is in fact a distinction between a good human will and a holy will: in contrast to God, even humans with the best wills feel inclinations whose satisfaction is possibly incompatible with the moral law. These inclinations constitute a "temptation" to action adverse to duty, and must therefore be suppressed or overcome. This danger of moral corruption that is an irrevocable part of the human condition is what Kant in the *Religionsschrift* calls "radical evil" (6: 32ff.): even the best human is radically evil (that is, evil "at the root"), since he can feel morally adverse inclinations (such as indolence, envy, selfishness, etc.), which, if they do not prevent him from following the moral law, at least make it more difficult. As we have already seen, however, in the case of a holy will all impulses are from the outset and necessarily in harmony with the demands of morality so that there is no temptation that would have to be overcome: in the case of a "holy (superhuman) being [...] no hindering impulses would impede the law of its will," and it "would thus gladly do everything in conformity with the law" (6: 405).

A good will is thus one that does what is morally required not simply accidentally, but out of firm resolution and lasting principle; a good will is holy if it does not have to overcome any inner resistance

(in the form of inclinations) to do so, but rather does the good, under any conceivable circumstances, "gladly." (To what extent holiness of the will can be demanded of humans, as Kant claims in the *KpV* and the *MdS*, is an additional problem that cannot be treated here; cf. on this Allison 1990, ch. 9).

### 5. *The concept of the imperative and the division of the imperatives* (4: 413–417)

"The representation of an objective principle, insofar as it necessitates a will, is called a command (of reason), and the formula of that command is called an imperative" (4: 413). According to this definition, an imperative is the linguistic expression ("formula") of a "command of reason," which itself is the "representation" of an objective principle or law that "necessitates" the will. (Kant does not always hold himself to this usage, though, and often uses the terms "imperative," "command," "law," and "objective principle" synonymously). "All imperatives," as Kant continues, "are expressed by an *ought*"—or, more precisely, *can* be expressed by an "ought," since "Act in such a way that the maxim of your action could be a universal law" is an imperative in Kant's sense, even though it is not "expressed by an ought." It, however, could easily be reformulated accordingly ("You ought to act in such a way, so that the maxim of your action ...").

The "necessitation" of the will that is expressed in the imperative is that specific rational form of normative motivation which consists in the awareness that one ought to do what is rational to do. Kant does not try to ground the connection between reason and normativity further: he presupposes here that a being that is endowed with reason but not always rational *ought* to be rational (cf. Kant's discussion of the usefulness of reason at the beginning of *GMS* I: 4: 395/6). He does, however, explicate the normative and evaluative dimension of reason a bit later by highlighting the connection between "ought" and "good": imperatives say "that to do or to refrain from something would be good [...]. Practical good, however [in contrast to the "pleasurable"; M. W.], is that which determines the will by means of representations of reason, hence not by subjective but objective causes, that is, from grounds that are valid for every rational being as such" (4: 413; cf. 414). According to Kant, every imperative implies a value judgment, but one that is rationally justifiable and thus claims to hold for all rational beings. (Here we are reminded of the famous thesis of

the *KpV* concerning the priority of the moral law over the concept of the good; cf. 5: 57–71.)

Imperatives, however, do not hold for all rational beings, because, as we have already seen, a holy will is not subject to any normative "necessitation": "the 'ought' is out of place here," for Kant, "because *volition* is already of itself necessarily in accord with the law" (4: 414). Kant is famous for the thesis that ought implies can: someone who *ought* to do A, also *can* do A (cf. 5: 30). In contrast, his claim here is: *ought* implies being able to do *otherwise*. Someone who ought to do A, can also *not do* A. (It does, for example, seem as senseless to command someone to be identical with himself, as it is to command him to *not* be identical with himself; he cannot stop doing the first, and he cannot do the second. In both cases the "ought" is "out of place.")

Kant's view can be summarized by saying that imperatives are *rational directives* (or directives of reason). Their normative binding force is based on the fact that it is *rational* to follow them. Their prescriptive character is based on the fact that they are directed toward beings that can act rationally, but do not necessarily always do so. Kant continues: "All imperatives command either hypothetically or categorically. The former represent the practical necessity of a possible action as a means to achieving something else that one wants (or that it is at least possible for one to want). The categorical imperative would be one that represented an action as objectively necessary in itself, without reference to another end" (4:414). On the basis of the analytic relationship between "ought" and "good," Kant can express the distinction between hypothetical and categorical imperatives in the following way as well: "Now, if the action would be good merely as a means to something else, the imperative is hypothetical; if the action is represented as in itself good [...], then it is categorical" (4:414).

The distinction between hypothetical and categorical imperatives thus has an impact on the question of why it is good and rational to act in the way commanded by the imperative. This distinction does not necessarily show in the linguistic form of the imperative: "Invest your money cautiously!" is, viewed linguistically, a categorical sentence but a hypothetical imperative, for a cautious investment is obviously good "merely as a means" and is not sought after for its own sake. "If you owe money, you ought to pay it back" is linguistically a conditional sentence but nevertheless a categorical imperative, because it represents an action "as objectively necessary in itself, without reference to



another end" (on the problem of the distinction between hypothetical and categorical imperatives cf. Brinkmann 2003, 19–51).

As mentioned above, Kant divides hypothetical imperatives further into "rules of skill" and "counsels of prudence" and thus arrives at a tripartite division of the imperatives:

(1) *Rules of skill* prescribe an action under the condition of an end concerning which it is left open whether one actually follows the end and whether it is rational to do so: "If you want A, then you ought to do B (as a necessary means thereto)!" Nothing changes in this hypothetical structure when the condition is met and one thus actually pursues the end in question: Kant does not provide for the derivation, using *modus ponens*, of a *categorical* imperative, "You ought to do B," from the hypothetical imperative "If you want A, you ought to do B" and the empirically-contingent observation "You want A." The imperative thus derived, "You ought to do B," remains a hypothetical one, as its validity depends upon the condition that the person in question wants A.

(2) For this very reason *counsels of prudence* are also merely hypothetical imperatives. They recommend an action as a means to an end that the addressee of the imperative actually has: "Because you want A, you ought to do B (as a necessary means thereto)!" Because imperatives are expressions of principles of reason and thus have *universal* validity, the only end in question here is one that all rational beings actually pursue (insofar as, as Kant notes, "imperatives apply to them as dependent beings," 4: 415). The only such end is that of one's own "happiness": every person wants to be happy. (Kant defines happiness as a state in which "everything goes according to one's wish and will" (5: 124)). This pursuit of happiness is a natural fact – it is part of what constitutes humans as rational living beings; thus the "purpose of happiness" is something we have by a "natural necessity" (4: 415). It can thus be presupposed to exist in every human being. Nevertheless, this purpose cannot serve as a basis for a categorical imperative, for different people find happiness in very different things. No means, therefore, can be identified that would be actually necessary for every person in order to become happy: "Because you (like every rational natural being) want to be happy, you ought to live healthily," or, more concisely, "Live healthily!" is a rational counsel, but with some luck (*Glück, fortuna*) one can also be happy (*glücklich, beatus*) if one does not heed it.

(3) Only *moral* (categorical) *imperatives* hold without exception for all ("dependent," i. e., capable of being sensibly motivated) rational

beings. They do not concern the "material of the action" (thus do not depend on which ends one is pursuing), but concern rather "the form and the principle from which it [the action] itself follows" (4: 416). The "form" of the action is, as has already been shown in the first section of the *GMS*, the "conformity to law as such," which serves the will "as its principle" (4: 402). In this manner, Kant arrives at the categorical imperative in the *singular* ("never to act except in such a way that I could also will that my maxim should become a universal law," 4: 402; cf. 421). According to Kant, out of this universal categorical imperative, as their principle, "all imperatives of duty," i. e., all particular categorical imperatives such as the prohibition on lying, stealing, etc., "are derived" (cf. 4: 402). What unites these distinct categorical imperatives is the fact that one ought to act in the prescribed way not for an "actual or possible purpose," but simply because the action is good "in itself." – To prove that there actually are such actions (and thus also the categorical imperative that commands them) is the aim of the third section of the *GMS*.

## Literature

### Kant's writings

Kant's writings will be cited according to the pagination of *Kants gesammelte Schriften*, Akademie Ausgabe (Berlin: deGruyter, 1902–) (abbreviated as 'AA'). The *Critique of Pure Reason* will be cited according to the A/B pagination from the first and second editions. All textual references are to *The Cambridge Edition of the Works of Immanuel Kant* (Cambridge University Press, 1992–).

GMS *Grundlegung zur Metaphysik der Sitten*, AA, IV  
 KdU *Kritik der Urteilskraft*, AA, V  
 KpV *Kritik der praktischen Vernunft*, AA, V  
 KrV *Kritik der reinen Vernunft*, AA III, IV  
 MAN *Metaphysische Anfangsgründe der Naturwissenschaft*, AA, IV  
 MdSR *Metaphysik der Sitten, Rechtslehre*, AA, VI  
 MdST *Metaphysik der Sitten, Tugendlehre*, AA, VI  
 RGV *Die Religion innerhalb der Grenzen der blossen Vernunft*, AA, VI  
 TG *Träume eines Geistersehers*, AA, II

### Other works

Allison, Henry E. (1990): *Kant's Theory of Freedom*, Cambridge.  
 Beck, Lewis White (1974): *Kants "Kritik der praktischen Vernunft." Ein Kommentar*, München.

- Brinkmann, Walter (2003): *Praktische Notwendigkeit. Eine Formalisierung von Kants kategorischem Imperativ*, Paderborn.
- Gerhardt, Volker (1986): "Handlung als Verhältnis von Ursache und Wirkung. Zur Entwicklung des Handlungsbegriffs bei Kant", in G. Prauss (Hg.), *Handlungstheorie und Transzendentalphilosophie*, Frankfurt/M.
- Laberge, Pierre (1989): "La définition de la volonté comme faculté d'agir selon la représentation des lois", in O. Höffe (Hg.), *Grundlegung zur Metaphysik der Sitten. Ein kooperativer Kommentar*, Frankfurt/M., 83–96.
- Timmermann, Jens (2003): *Sittengesetz und Freiheit. Untersuchungen zu Immanuel Kants Theorie des freien Willens*, Berlin.
- Willaschek, Marcus (1992): *Praktische Vernunft. Handlungstheorie und Moralbegründung bei Kant*, Stuttgart/Weimar.
- Wood, Allen (2006): "The Good Without Limitation", in this volume.

Bernd Ludwig

## Kant's Hypothetical Imperatives (GMS, 417–419)

### I.

In German language (and also in many others) there is a grammatical trap, which usually snap shuts as soon as one starts talking about Kant's imperatives thoughtlessly. It goes like this:

'Who wants to play the piano must practice.' This is a true statement informing us of connections in the world. It is almost exemplary descriptive and would be wrong if for example playing the piano was like sneezing, which one does *not* have to practice to be able to do it. Put into second person, this statement reads 'If you want to play the piano you must practice.' This is another descriptive statement, particularly since it was generated by insertion into the former – a procedure which does not affect the descriptive character.<sup>1</sup> The latter can also be expressed differently: 'If you want to play the piano well, practice!' It is true that the tone might change slightly (maybe from an amicable hint to a parental admonition), but put aside the educational authority (and the possibly connected sanctions) nothing has changed but the *grammatical* form. The *grammatical* indicative 'you must practice' has been substituted by the *grammatical* imperative 'practice!', including an expression mark at the end of the sentence.

Now the mentioned trap is wide open, for expression marks also appear at the end of sentences which not only *contain* the grammatical form of an imperative but actually *are* imperatives ('Practice!'), i. e. in Kant's terminology: statements expressing an 'ought' and thus 'an objective necessitation [Nöthigung]' towards a certain action (see KpV, 20). This kind of statement one can address to oneself as well

<sup>1</sup> This kind of insertion is, for instance, used when trying to take into account the peculiarities of the addressee, like: 'If you want to play the piano well, YOU have to practice (in opposition to your advanced brother)'.