

From

A Companion to Kant

Graham Bird, ed.

Blackwell 2006

18

Kant's Critical Account of Freedom

ANDREWS REATH

I. Introduction

Kant's treatment of free will is one of the most intriguing, as well as most perplexing, elements of his philosophical system. His Critical account of free will is given mainly in the Resolution of the Third Antinomy of the *Critique of Pure Reason* (B 559–86) and the Analytic of the *Critique of Practical Reason*, but important components are found in the third section of the *Groundwork of the Metaphysics of Morals*, the *Metaphysics of Morals*, and Book I of *Religion within the Boundaries of Mere Reason*. It draws on and requires some understanding of several aspects of his system, but principally his transcendental idealism and its distinction between phenomena and noumena, or appearances and things in themselves, and his account of the moral law as a principle of autonomy that gives reasons for action that are independent of and take priority over reasons based on desire. The former is developed in the first *Critique*, and the latter in the *Groundwork* and the second *Critique*.

Kant defends what at first might seem to be an impossible position. Having established the principle of causality by arguing that it is a necessary condition of possible experience, Kant accepts a strict form of causal determinism for the natural world, according to which every event follows necessarily from prior events according to empirical laws. At the same time he argues that we are warranted in ascribing to ourselves a strong libertarian form of free will, according to which rational agents have the capacity to choose courses of action independently of determination by antecedent conditions, and that, accordingly, alternate possibilities are open to them at the time of action. In particular, in a situation where an agent fails to do what she ought to have done, she could have chosen what she had most reason to choose. Kant thinks that moral responsibility requires no less.

What allows Kant to avoid outright contradiction in accepting both strict causal determinism and libertarian free will is the distinction between phenomena and noumena, and it will be helpful to begin with some background on that distinction. Kant's transcendental idealism holds that space and time are the "forms of intuition"; they are not mind-independent features of reality, but structural features of experience specific to human cognition (see chapter 7 above). The properties of objects and events as they are given in space and time, including their causal relations to prior events, are not

properties of things as they are in themselves but of objects as they appear to us. By claiming that spatiotemporal properties are features of objects as they appear to us, transcendental idealism creates the possibility of thinking about objects under two aspects, both as phenomena and as noumena. Since knowledge requires grounding in intuition and things in themselves are not given in intuition, knowledge is limited to objects as they appear to us. But while there is no knowledge of noumena, there is still room for thought about noumena. For example, we can coherently entertain the thought of objects as they are in themselves. That is just the thought of ordinary objects apart from the conditions under which they appear to us – though for that reason, interactions between things in themselves (or between things in themselves and our minds) cannot be characterized. As we will see, Kant thinks that we can make warranted assertions about noumena if we have grounds for thinking of entities that are governed by laws different in kind from the causal laws governing the occurrence of events in space and time. Such assertions would be a way of conceiving of such entities, but are not *knowledge* claims. Noumena are entities that can be thought though not known, but warranted assertions about noumena are possible if they can be grounded in something other than intuition or the conditions of possible experience.

The distinction between phenomena and noumena opens up the possibility of viewing human actions under two different aspects or from two different standpoints. When we view human beings as phenomena, we treat their actions as events in the natural world that are causally determined by facts about their psychology and their circumstances. So regarded, actions are to be explained by a person's desires and interests, and by the psychological traits that Kant terms a person's "empirical character." Likewise a person's empirical character can be causally explained in terms of formative influences such as the person's environment, native temperament, and so on. But Kant argues that our possession of various rational capacities, including the capacity to guide our activity by various rational norms, warrants ascribing to ourselves the power to choose independently of determination by antecedent conditions. When we think of ourselves as exercising this kind of causality, namely free agency, we regard ourselves as noumena. A person's "intelligible character" is the set of fundamental principles and value priorities that guide an agent's choices. We think of the intelligible character as under a person's control, and it is by seeing a person's actions as manifestations of his intelligible character that we trace them back to his activity. By this route, Kant's transcendental idealism permits him to argue that strict determinism in the natural world does not rule out the possibility of free will. As phenomena, human actions are causally determined by antecedent conditions that extend back in time. But when we think of ourselves as noumena, we ascribe free agency to ourselves, and the same actions may be thought of as flowing from an agent's free choices.

The intuitive appeal of Kant's theory is that we do adopt different standpoints on human action, both of which seem deeply rooted in our thought. We believe that human actions are parts of the natural world and can be explained by tracing them back to prior conditions. In this context, understanding takes the form of seeing how an action results from a person's psychological states and character traits, which in turn result from past influences according to psychological laws. This is sometimes called the "third person" or "theoretical" standpoint. But as rational agents who think and act, we take

up a very different standpoint towards ourselves, a "first person" or "practical" standpoint. As thinkers and agents we face the practical or deliberative tasks of determining what we have reason to believe or what we have reason to choose in various situations. Engaging in these practical tasks appears to carry certain presuppositions. We are aware of the capacity to guide our activity and choices by a wide range of rational principles, such as norms that govern inference and the formation of belief or principles of rational choice that determine how we ought to act. We suppose that our choices and decisions are settled by our judgments of what we have reason to believe or to do. In that sense they are up to us. This is a capacity for self-determination that is independent of certain kinds of causal determination, because the idea of a reason introduces a *different* kind of connecting ground from that seen in causal connections between events. The connections between a set of premises and the conclusion that follows, between a body of evidence and a belief, between a set of justifying reasons and a choice, and so on, are normative rather than causal connections. Reasoning and judgment thus appear to involve a capacity for self-determination and, when they guide choice, a capacity to initiate courses of action that is not governed by causal laws.

Arguably this is how we conceive of ourselves when we engage in reasoning and deliberation, and how we conceive of others when we regard them as responsible agents. But can this conception of ourselves be sustained in the face of causal determinism? Kant aims to show that it can be. His account of free will unfolds in different stages. In the first *Critique*, he argues that libertarian free will is not ruled out even if one accepts strict determinism, as he does. He tries to establish the possibility of free will through the distinction between phenomena and noumena and the idea that normative principles that are essential to rational activity are different in kind from causal laws. Some passages in the first *Critique* and in the *Groundwork* appear to suggest that our rational capacities of understanding and theoretical reason warrant ascribing freedom to ourselves. But Kant's final view, seen in the second *Critique*, is that the reality of freedom is established through our moral consciousness.

Sections II and III of this essay will concentrate on Kant's claims in the first *Critique* that transcendental idealism creates room for the possibility of free will. The final section touches briefly on Kant's claim in the second *Critique* that our consciousness of the authority of the moral law establishes its reality. Kant's attempt to establish the possibility of free will is often dismissed because it is thought to rely on a very strong and implausible metaphysical conception – that the acting self exists in a noumenal realm outside of time, free from causal determination, and that its free choices are the ground of its actions in the phenomenal world. There is certainly textual material that suggests such a conception, but I develop a different reading. I treat the phenomenal and the noumenal as two different standpoints that we adopt toward action in different contexts. To regard human beings as noumena is simply to take up the practical standpoint and regard ourselves as agents who act for reasons. Likewise Kant's view that reason, or the noumenal self, is not subject to the conditions of time may be understood in terms of the idea that rational activity is guided by normative principles, which are essentially different in kind from empirical laws that determine how events follow from temporally prior conditions.

II. Transcendental Freedom and Practical Freedom

The Resolution of the Third Antinomy has three sections. The first distinguishes different notions of freedom that shape Kant's understanding of the free will problem (B 560–5). The second gives what he terms “a silhouette of a solution” (B 570), that sketches how transcendental idealism might preserve the notion of free agency required for moral responsibility (B 566–9). The third is a “Clarification” of how free agency in human beings can be consistent with causal determinism (B 570–85). In this section, I discuss Kant's distinction between “transcendental freedom” and “practical freedom,” and in the next section outline his account in the first *Critique* of the possibility of freedom (supplemented with material from other important texts).

“Transcendental freedom” is the cosmological idea of a spontaneous first cause. Kant defines it as a form of causality “through which something happens without its cause being further determined by another previous cause, i.e., an absolute causal spontaneity beginning from itself” (B 474). Or as he says elsewhere, it is “the faculty of beginning a state from itself, the causality of which does not stand in turn under another cause determining it in time in accordance with the law of nature” (B 561). It is the capacity to initiate a series of events without being determined by temporally prior causes. It is important to note that transcendental freedom is a form of causality – a way of making things happen – though one that differs from natural causality. In contemporary terms, natural causation is event causation. Transcendental freedom, however, is not a form of event causation, since the activity of a transcendently free cause is not determined by temporally prior events or conditions. However, to say that transcendental freedom is not determined by prior causes does not imply that it is undetermined. Throughout the Critical philosophy Kant makes clear that any cause must be governed by some law which connects the cause or its activity to its effects, and the same holds for freedom as a causality “of a special kind” (G, 4.446). Transcendental freedom must therefore be a form of causality that operates on principles *different in kind* from empirical laws of nature, though he does not specify what these principles are when he first defines this idea.

While transcendental freedom is an abstract idea that arises in cosmological contexts, “practical freedom” is the form of free agency that seems to be supported by our first-person experience of ourselves as agents. Kant writes: “Freedom in the practical sense is the independence of the power of choice from *necessitation* by impulses of sensibility” (B 562). The human power of choice is “pathologically affected” but not “pathologically necessitated.” To explain, we are moved by sensible desires and interests, but they do not directly cause our actions. The human power of choice appears to be free “because sensibility does not render its action necessary, but in the human being there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses” (B 562). Elsewhere Kant says that a power of choice “which can be determined independently of sensory impulses, thus through motives that can only be represented through reason, is called *free choice* (arbitrium liberum), and everything that is connected with this . . . is called *practical*” (B 830; cf. also *MM*, 6.213). In these passages Kant defines practical freedom as the power to act on principles of reason, independently

of causal determination by sensible motives. It is a power of self-determination, because it involves the power to determine oneself to action through one's application of rational principles, both principles of instrumental and prudential rationality and moral principles.

The idea that practical freedom involves independence of determination by sensible motives contains different elements worth distinguishing. In these passages, Kant certainly says that sensible desires do not directly cause our actions, and that we can set aside even powerful desires when they conflict with our judgments (both prudential and moral) of what we ought to do. This dimension of practical freedom is developed further in an important passage from the *Religion* that has come to be known, following Henry Allison (1990), as "The Incorporation Thesis":

The freedom of the power of choice has this characteristic, entirely peculiar to it, that it cannot be determined to action though any incentive, *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom). (RBR, 6.23–4; translations from Kant 1996 and 1998)

It is a feature of free agency that a desire or motive can influence choice and lead to action only when the agent "incorporates it into his maxim" – roughly, when the agent takes the motive to give him a good reason for action and adopts the subjective principle of acting on that motive. The "incorporation" of an incentive is an agent's spontaneous act of taking it to be a good reason, judging that it is worth acting on, or in some way endorsing it. It is important to note that the Incorporation Thesis is a claim about the influence of *any* incentive on choice, including rational motives and considerations. It claims that the influence of any incentive on choice always passes through a spontaneous act or judgment on the part of the agent; roughly, motives of all sorts lead to action only when we choose to act on them. But Kant's discussions also suggest that we are not bound to take our sensible desires as reason-giving. We can, for example, judge that acting on a certain desire would be injurious, unworthy, or morally wrong, and that it provides no reason for action. As is clear from his moral theory, Kant believes that we have the capacity to act from reasons that make no reference to our sensible desires.

Implicit in the Incorporation Thesis is a further sense of "independence of determination by sensibility" that is central to Kant's account of free will. The very idea of rational judgment (e.g., a judgment about reasons) rules out causal influence. Kant claims that a rational agent necessarily "acts under the idea of freedom," and says that reason cannot "receive direction from any other quarter" since the resulting judgment would then be determined not by "reason" but by an "impulse" (G, 4.448). Judgments understood as caused according to psychological laws cannot be regarded as a rational. More generally rational activity is "independent of determination by sensibility" in the sense that it does not operate according to empirical laws, but is guided by normative principles, which are essentially different in kind.

Another important feature of practical freedom appears in Kant's claim that practical freedom

presupposes that although something has not happened, it nevertheless *ought* to have happened, and its cause in appearance was thus not so determining that there is not a causality in our power of choice such that, independently of those natural causes and even opposed to their influence, it might produce something determined in the temporal order in accord with empirical laws, and hence begin a series of occurrences *entirely from itself*. (B 562; see also B 583)

Often we fail to act as we ought to. For example, I might be inclined to act in a certain way, decide that it would be imprudent or wrong, but nonetheless choose that action. It is part of our ordinary conception of our agency that we can do what we judge that we ought to do, or what we have reason to do. That means that in situations like these, we believe that we could have acted otherwise – that alternative possibilities were open to us. In this example I acted on a sensible motive, contrary to my judgment of what I had most reason to do, but that motive “was not so determining” as to close off the possibility of having acted differently. I could have set that motive aside and done what I had most reason to do.

In sum, the practical freedom that we ascribe to ourselves is a capacity for self-determination, or a capacity to initiate action through one’s judgments about reasons. It is a form of agent causation that is guided by an agent’s reasoning. There are different senses in which this capacity is not sensibly determined. Motives do not directly cause action, but rather influence choice through a spontaneous act or judgment by the agent; and we are not bound to take sensible motives as reasons, but can set them aside or oppose them. Further, practical freedom involves the possibility of acting otherwise, since in cases where we act against reason, we suppose that we could have done what we judged we had reason to do. Finally, it is independent of sensible determination in the sense that it does not operate according to empirical causal laws, but rather according to normative principles. Since this practical freedom appears to be a power to initiate actions that is not determined by temporally prior causes, it seems to be an instance of transcendental freedom, and therein lies the problem. Since nature is a deterministic system, there is no room in nature for transcendental freedom. That raises the question whether the practical freedom that we ascribe to ourselves is illusory. Since “the practical concept of freedom is grounded” on the transcendental idea, “the abolition of transcendental freedom would also simultaneously eliminate all practical freedom” (B 561, B 562).

Kant rejects compatibilist solutions to the free will problem as “a wretched subterfuge” (CPrR, 5.96). Compatibilists adopt what Kant calls a “comparative concept of freedom.” They hold that human actions are free when “caused from within, by representations produced by our own powers, whereby desires are evoked on occasion of circumstances and hence actions are produced at our own discretion” (CPrR, 5.96). In other words, actions are free when caused by an agent’s internal psychological states, such as beliefs, desires, and intentions. Kant argues that this conception of freedom falls short of the practical concept outlined above. The psychological states that compatibilists identify as causes of action are still events that follow causally from prior events, in this case according to psychological laws. Since psychological causation is still natural causation, the compatibilist solution does not give the agent any real control at the time of action. Their conception of freedom is “nothing better than the freedom of the turnspit, which, when once it is wound up, also accomplishes its movements by itself” (CPrR, 5.97).

How then does Kant establish that the free will required for agency and responsibility is possible? Since Kant holds that the law of causality applies to all events in the natural world, and since human actions are such events, free agency is possible only if actions may be viewed under different aspects as the effects of both natural causation and free agency. This is not an option for the transcendental realist: if spatiotemporal determinations are properties of things in themselves, then descriptions of events in terms of such properties and deterministic causal laws are exhaustive. But the distinction between appearances, phenomena, and noumena may enable the transcendental idealist to view human action in these two different ways.

III. The Possibility of Freedom of the Will

The following passage articulates a principle that guides Kant's solution to the free will problem:

I call *intelligible* that in an object of sense which is not itself appearance. Accordingly, if that which must be regarded as an appearance in the world of sense has in itself a faculty which is not an object of intuition through which it can be the cause of appearances, then one can consider the *causality* of this being in two aspects, as *intelligible* in its *action* as a thing in itself, and as *sensible* in the *effects* of that action as an appearance in the world of sense. (B 566)

Here Kant claims that if an entity in the natural world has a capacity to bring about changes in the world that is "intelligible" and "not an object of intuition," then its causal powers and relations may be considered under two aspects. As noumenon it would have a capacity for activity that is "intelligible" – presumably a capacity to guide its actions by normative principles – while as phenomenon its actions would be causally determined events in the natural world. Kant goes on to claim that such an entity would have both an empirical character, which determines how its actions follow from temporally prior conditions, and an intelligible character, which guides its rational activity as noumenon. In its intelligible character it "would not stand under the conditions of time" (B 567), which is to say that its actions would not follow causally from temporally prior conditions according to empirical laws. Since it has the power to initiate actions without being determined by temporally prior conditions, this entity would satisfy the definition of transcendental freedom.

Relying on this principle, Kant constructs roughly the following argument in the Clarification:

- 1) Since the principle of natural causality holds without exception of all events, free will is possible only if human actions can be regarded, under different aspects, as effects of both natural causation and free agency.
- 2) If an entity in the natural world has a capacity to effect changes that is intelligible and not an object of intuition, its causal powers may be viewed under two aspects and will have both an empirical character and an intelligible character. As phenomenon its actions would be explainable in terms of its empirical character. But

as noumenon, its activity will be guided by intelligible (i.e. normative) principles, according to its intelligible character.

- 3) In its intelligible character, this entity will have a capacity to initiate actions independently of determination by temporally prior conditions, and thus will be transcendently free.
- 4) Human beings possess rational capacities that are intelligible and not objects of intuition. These appear to include the capacity to guide their choices and actions by rational principles – that is, it appears that “reason has causality” (B 575) and is a power to effect changes in the world. If so, human beings have a causal power that is intelligible and they satisfy the principle in (2) above: their causal powers and relations may be viewed under two aspects and would have both an empirical and an intelligible character.
- 5) Assuming that reason is a causal power that can lead to action (as in (4) above), human beings are transcendently free in their intelligible character. They have a capacity to initiate actions through the exercise of their rational capacities, even though these same actions can also be given empirical explanations in terms of psychological facts about the agent.

Initially this argument raises more questions than it answers, and there is much that needs explaining. I shall focus on three general questions: (a) why does the possession of rational capacities require that we view the causal powers of human beings under different aspects and what does it mean to say that rational capacities are “intelligible” and “not objects of intuition”? (b) What is the relation of the empirical character to the intelligible character? And (c) what does it mean to say that “the acting subject, in its intelligible character, would not stand under any conditions of time,” and how would that support the claim that human beings, considered as noumena, are transcendently free?

To begin with the first question, Kant draws on the fundamental insight that rational activity is guided by normative principles, which are different in kind from causal laws, and that it cannot be understood in terms of empirical causal principles. Accordingly, when we ascribe rational capacities to ourselves, as we do in taking up the practical perspective, we regard our activity differently than when we think of human actions as events in the natural world subject to causal determination; we view our powers and activity “under a different aspect” because we understand it in terms of normative principles. To see this, consider what goes on in a chain of reasoning when a person draws a conclusion from an argument. The person draws normative connections between the steps of the argument, judging that one step follows from another and that the conclusion follows from the entire argument. She comes to accept the conclusion because she judges that the argument provides rational grounds for accepting it. In valid reasoning, the conclusion does not follow from its rational ground (i.e., from the supporting argument) in the way that an effect follows from its cause according to empirical laws. In this case, the conclusion follows from the argument via the relevant principles of inference, and her judgment that it follows is the reason why she accepts the conclusion. The connections here – the logical or rational connection between premises and conclusion, and the connection between the person’s going through the argument and her accepting the conclusion – are normative not causal, because they are given by norms of inference and rationality.

The same things can be said about an agent who judges that she has a sufficient reason to choose a particular action, forms the relevant intention and acts accordingly; in Kant's terms, she "incorporates" an incentive into a maxim of action. Her judgment about what she has reason to do is the rational ground of her action, and her action does not follow from its rational ground in the way that an effect follows from its cause according to empirical laws. In this case, the action follows from the judgment through the norm of rationality to the effect that someone who judges that he has most overall reason to perform a certain action in a given situation rationally ought to perform that action. (Someone who makes this judgment but for no good reason fails to perform the action displays a form of irrationality.) The agent is exercising a form of causality since he acts and brings about changes in the world on the basis of his rational judgment. But it is not empirical causality, since it is not governed by empirical causal laws. It is the kind of causality involved in rational agency.

If we do try to understand a rational process in terms of empirical causal laws, we lose the sense of it as rational. For example, consider trying to understand the reasoning in the above two cases as a causal process in which one psychological state or set of states (rehearsing steps of the argument, the judgment about what one has reason to do) causes subsequent psychological states (accepting the conclusion, the motivation or intention to act) according to psychological laws. We would be thinking of the relevant psychological states simply as events in the person that follow causally from prior events. But once we think of the process in these terms, we lose the sense that there is an *agent* who is *drawing normative connections* between the relevant items (from premises to conclusion, or from a judgment about reasons to action). These notions need to be in play if we are to think of the process as rational activity.

Normative principles tell us what we ought to think or do, and Kant says that "ought expresses a species of necessity and a connection with grounds which does not occur anywhere else in the whole of nature" (B 575). In addition to providing a ground of action that differs from empirical causal grounds, they introduce a possibility that does not apply to events that are part of the natural causal order, namely the possibility of actions other than those that actually occurred. We assume that rational agents can do what they judge they ought to do. Thus when agents act contrary to reason we assume that they could have acted otherwise. They could have done what they had reason to do.

We can now explain why rational capacities are "intelligible" and "not given in intuition." A capacity is "intelligible" if it is conceived through concepts and principles that originate in understanding or reason. Rational activities such as reasoning and judgment are intelligible in this sense because in order to understand them in their distinctive character as rationally guided, we must bring them under normative concepts and principles, as explained above, and such concepts and principles have an a priori origin in understanding and reason. Neither our possession of rational capacities nor particular exercises of rationality are given in sensible intuition (e.g., as items or events in the natural world) because they are not objects of observation in the normal sense, to be understood by bringing them under empirical causal principles. One feature of Kant's epistemology is that objects of sensible intuition are brought to the unity of consciousness by bringing them under empirical concepts whose form is given by the categories of the understanding, and any event given in intuition must follow from

temporally prior conditions according to empirical laws. For something to be a possible object of intuition, it must be amenable to this form of understanding. But as we have seen, rational activity cannot be understood in this way, because it loses its character as rational if we try to bring it under empirical causal laws. In that specific sense, neither rational capacities nor instances of their exercise (such as acts of reasoning, judgment, choice, and so on) are given in intuition.

Turn now to the distinction between empirical and intelligible character, which is simply a consequence of the idea that our causal powers may be viewed under two aspects. Kant writes that every cause "must have a character, i.e., a law of its causality, without which it would not be a cause" (B 567). The "character" of a form of causality would be the principle according to which it operates. A person's empirical character would be the principles that connect his actions with temporally prior conditions, while the intelligible character would be the set of basic (normative) principles that determine how one exercises one's practical reason. Starting with the latter, two features of the intelligible character are worth noting. First, one's intelligible character is the set of basic principles, value commitments and priorities, and maxims that guide one's choices by determining what one sees reason to do in various circumstances. Second, it will be these basic principles, understood as originating in that person's rational agency – that is, as principles and values that one has in some sense adopted or endorsed, and for which one is responsible. The intelligible character is intended to capture the notion of a person's moral character since it includes a person's basic principles and values, thought of as adopted or endorsed by the person.

The empirical character will have two corresponding features. First, Kant understands a person's empirical character as a set of rules or laws that specify how his actions follow from temporally prior conditions and that may be inferred from the person's observed actions (cf. B 567, B 577). Since it involves rules or laws, it includes facts about the person's psychology such as standing desires and dispositions or motivational tendencies that determine, and may be cited to explain, how the person acts in various circumstances. Second, it treats these standing dispositions and motivational tendencies as themselves subject to empirical causal explanation. Presumably the dispositions that make up the person's empirical character follow from such temporally prior conditions as the person's upbringing and social environment, native temperament, past experiences, and other formative influences, according to psychological laws or generalizations. In this respect, empirical character is a naturalistic explanatory notion that locates human actions in a temporally extended causal process.

Matters are complicated somewhat by the fact that a person's reason exhibits an empirical character. Kant writes:

every human being has an empirical character for his power of choice, which is nothing other than a certain causality of his reason, insofar as in its effects in appearance this reason exhibits a rule, in accordance with which one could derive the rational grounds and the actions themselves . . . and estimate the subjective principles of the power of choice. (B 577)

The implication here is that a person's empirical character includes a person's reasons (the "rational grounds" of action) and subjective principles, in other words, the maxims from which the person acts. If so, it is natural to include such items as a person's

observed value commitments, tendencies to take certain considerations as reason-giving, or even tendencies to reason in certain ways. One might think that these are normative dispositions that belong to the intelligible character, but the puzzle can be resolved in this way. We may assume that the empirical character includes basic principles, value priorities, tendencies to reason in certain ways, and so on, viewed simply as psychological facts about the person that can be inferred from observing a person's choices, and that are susceptible to empirical explanation in terms of past influence. The empirical character takes a purely descriptive approach to these dispositions, focusing on the principles that the person actually accepts and treating them simply as dispositions and motivational tendencies that can be cited in empirical explanations of actions.

Kant holds that the intelligible character may be regarded as the ground of the empirical character, and this idea is crucial to making sense of the idea that human actions may be viewed as resulting both from natural causation and free agency. As phenomena, actions are causally explainable in terms of the dispositions and motivational tendencies that comprise the empirical character. But Kant says that it is possible that this "empirical causality" has an intelligible ground in the person's intelligible character (B 572). What is possible, in other words, is that the dispositions observable in a person's actions are grounded in the person's basic principles and value commitments. Since commitments to principles, judgments about reasons, acts of "incorporating incentives" into maxims, and so on, are not given in intuition (in their character as rational), they are not items in our empirical understanding of action, but are introduced only when we ascribe rational capacities to human beings.

Two ideas may be at work in the thought that the intelligible character is the ground of the empirical character. First, the motivational tendencies that go into the empirical character may reflect principles and value commitments that are their ground. Someone who thinks that honesty is important unless it requires personal sacrifice will display a certain pattern of motive, feeling, and action. Likewise someone who judges that on balance he ought to be honest in a particular situation will be motivated to act honestly. When their actions are regarded as phenomena, the selective disposition to act honestly, of the motivation to be honest in a particular situation, will figure in a causal explanation of the action, and, as psychological states of the agent, will be explainable in terms of antecedent conditions. But these empirically given motivations also reflect the agent's acceptance of a principle, or his assessment of the reasons in those circumstances; thus the motivations that appear as the empirical cause have an "intelligible ground" in the rational activity of the agent. Second, according to Kant's Incorporation Thesis, even desires or motivational tendencies that are best understood simply as causal products of past influence only lead to choice when the agent incorporates them into a maxim; their efficacy as motives is grounded in an act of the agent. Here consider a person in a situation that elicits a strong inclination to act dishonestly who acts on that inclination; assume further that the inclination is explainable in terms of a disposition that can be traced to facts about his upbringing. This inclination and its causal antecedents will figure in an empirical explanation of the action as phenomenon. But when we view the person as an agent who acts for reasons, we introduce an additional item into our understanding of the action that does not show up, as it were, when the action is regarded as phenomenon – namely, the "incorporation" of

the incentive of dishonesty into a maxim of action. It is through a spontaneous act of the agent that the inclination to dishonesty becomes the operative incentive that leads to action. That is, the motivational state that we might identify as an empirical cause of the action becomes efficacious through a spontaneous act, and thus its "empirical causality" is grounded in the agent's intelligible character.

The thesis that the intelligible character is the ground of the empirical character expands our conception of the grounds of human action. When we limit ourselves to the material given in intuition, empirical explanations are adequate for certain purposes and complete as far as they go. Successful empirical explanations provide an understanding of actions by showing how they follow from temporally prior conditions. But taking up the practical perspective and thinking of ourselves as rational agents introduces additional items such as commitment to principles, reasoning, judgment – in a word, elements of normative guidance – that do not show up in the empirical standpoint since they are not given in intuition. These elements enlarge our conception of the causality underlying human action. As we have seen, they permit the thought that an intelligible act (such as the adoption of a maxim, the application of a principle, or a judgment about reasons) is the ultimate ground of the action because it is the ground of the motivational state that appears as the empirical cause of the action – that is, that empirical causal factors have a further ground in the activity of an agent that enables us to see the agent as the source of his or her actions. Furthermore, because rational choice is guided by "ought" judgments, the elements of normative guidance introduce the possibility of acting otherwise in circumstances in which an agent acts contrary to reason.

Once such concepts are on board, it is clear that empirical explanations do not complete the story of human action because they omit one of its essential features, namely the rational activity of the agent. Presumably Kant needs to hold that the empirical understanding of action is incomplete in order to deal with a looming problem. He wants to say that actions as phenomena are causally determined and follow with necessity. But as noumena they are the results of free agency, which involves the possibility of acting otherwise. But isn't it just an outright contradiction to hold both that actions follow with necessity and that agents have the possibility of acting otherwise? The fact that the possibility of acting otherwise is only introduced by taking up the practical perspective and thinking of ourselves as rational agents may help with this problem. When we view actions as phenomena they do indeed follow from prior events that are, based on what is given in intuition, sufficient to bring them about. From that perspective, we cannot give any content to the idea that actions could have occurred differently. But thinking of ourselves as rational agents introduces further items into our conception of the causality underlying action that are not part of our understanding of actions as phenomena. Among other things, it allows us to say that the agent could have chosen differently (in situations where he or she chose contrary to reason). The possibility of acting otherwise is only introduced when we think of actions as noumena. But once we have a way to give content to that possibility, we see that our understanding of action based on what is given in sensible intuition is limited, though empirical explanations remain adequate for certain purposes.

One of the most perplexing features of Kant's account of free will is the idea that reason in its intelligible character does not "stand under any conditions of time"

(B 567). This idea is part of his argument that human beings, considered in terms of their intelligible character, act freely, and is seen in the following representative passages:

Pure reason, as a merely intelligible faculty, is not subject to the form of time, and hence not subject to the conditions of the temporal sequence. The causality of reason in the intelligible character *does not arise* or start working at a certain time in producing an effect. For then it would itself be subject to the natural law of appearances, to the extent that this law determines causal series in time, and its causality would then be nature and not freedom. (B 579–80)

But of reason one cannot say that before the state in which it determines the power of choice, another state precedes in which this state itself is determined. For since reason itself is not an appearance and is not subject to any conditions of sensibility, no temporal sequence takes place in it even as to its causality, and thus the dynamical law of nature, which determines the temporal sequence according to rules, cannot be applied to it. (B 581)

In regard to the intelligible character . . . no *before* or *after* applies, and every action, irrespective of the temporal relation in which it stands to other appearances, is the immediate effect of the intelligible character of pure reason; reason therefore acts freely, without being determined dynamically by external or internal grounds temporally preceding it in the chain of natural causes, and this freedom of reason can not only be regarded negatively, as independence from empirical conditions . . . but also indicated positively by a faculty of beginning a series of occurrences from itself . . . (B 581)

According to a common reading of such passages, Kant secures free will by placing the noumenal self outside of time, thereby freeing it from causal determination. Actions as phenomena are then somehow grounded in the atemporal and therefore free choices of the noumenal self. This conception is indeed suggested by these and other passages. But if Kant's resolution of the free will problem depends on such strong and, to most philosophers, dubious, metaphysical assumptions, its philosophical interest will be limited. We do better to look for a more innocuous reading of such passages, and I shall. (Allen Wood (1984) defends the "strong," "timeless agency" reading, while Henry Allison (1990) develops a metaphysically more "innocent" interpretation.)

The idea that reason, or agents considered in terms of their intelligible character, are not subject to temporal conditions may be understood through what we earlier identified as the distinctive feature of rational activity – that it is guided by normative principles, which are different in kind from empirical causal laws. We can put it this way: to say that reason is not subject to temporal conditions is just to say that the relation "is the rational ground of" is not the temporal relation "is the empirical cause of," since that relation is understood through normative principles rather than causal laws. Thus when Kant says that "no temporal sequence takes place in [reason] even as to its causality," he need not deny that actual reasoning takes place in time. The point is rather that the sequence in a rational process is not a causal sequence in which one state or event arises from preceding events according to empirical laws. The rational grounds of an action (or of a judgment, or an the agent's drawing a conclusion) are not its "antecedent conditions" in the way that an empirical cause is an antecedent condition of an effect – because to be an antecedent condition in that sense just means one from which the subsequent occurrence follows according to empirical laws. And

actions (or judgments, conclusions, and so on) do not follow from their rational grounds in that way since their connection to their grounds is normative. Likewise, in saying that "the causality of reason in the intelligible character does not arise or start working at a certain time in producing an effect," he means that reasoning is not understood in its character as rational or normatively guided through the kinds of laws that govern the occurrence of events in time.

We saw in the last section that Kant defines transcendental freedom as the capacity to initiate a series of events without being determined by temporally prior causes, and he claims that practical freedom – the form of free agency implicit in the practical perspective – presupposes transcendental freedom. Our ascription of rational capacities to ourselves and the conception of an intelligible character support the idea that human agents are transcendently free by supplying both the negative and the positive components of transcendental freedom. First, since rational activity is normatively governed, the elements of a rational process do not follow from temporally prior states according to empirical laws. ("[T]his freedom of reason can not only be regarded negatively, as independence from empirical conditions . . .") Second, normative guidance introduces the idea of spontaneous acts of an agent, such as an agent's judging that a conclusion follows from an argument, judging that there is reason to act in a certain way, applying a principle to a situation, and so on. When acts of this sort lead to action, we would have an instance of an agent initiating an event in the world. This is the positive component of transcendental freedom (The "freedom of reason" is "also indicated positively by a faculty of beginning a series of occurrences from itself . . .") Thus, in virtue of having a capacity to initiate events without being determined by temporally prior causes, human beings would satisfy the definition of transcendental freedom.

Kant gives an example of a person who tells a malicious lie that illustrates how we move between the empirical character and the intelligible character. Kant supposes that we can give an empirical explanation of the action in terms of "the sources of the person's empirical character," such as his upbringing, social environment, and native temperament. Such factors and the resulting motivational state would be the "occasioning causes" of the action as phenomenon. However, we might still blame the agent, and if so, we regard the agent's "reason as a cause that, regardless of all the empirical conditions just named, could have and ought to have determined the conduct of the person to be other than it is" (B 583). If we blame the agent, we think of his agency in a certain way. For example, though the agent may have been strongly inclined to lie (due to past formative influences), we suppose that this incentive leads to action through the agent's judgment that the benefits achievable by lying were a sufficient reason to lie. In that way we trace the action back to the agent. Further, we suppose that the agent could have refrained from lying. He had reason not to lie, had access to that reason, and could have done what he judged he ought to have done. It is worth noting that Kant's claim here is conditional: *if* we blame the agent despite our belief that the action has empirical causes, then we are supposing that the agent initiated the action without being determined by antecedent circumstances, and that he could have chosen as he ought to have. The distinction between phenomena and noumena, between the empirical and the intelligible character, is intended to show how we can ascribe free agency even though we believe that the action is empirically caused.

IV. The Reality of Freedom of the Will

The Resolution of the Third Antinomy only shows that the possibility of free will is not ruled out by causal determinism. To establish that human beings actually have free will, Kant needs to show that we can be motivated to act by principles of reason, in which case, reason is a "form of causality." We would then be entitled to employ the conception of intelligible character and ascribe transcendental freedom to ourselves – entitled in the sense that our awareness of these rational capacities is a ground for ascribing such capacities and doing so is consistent with causal determinism.

The argument that we can be motivated to act by reason in the requisite way comes in Kant's moral theory. In *Groundwork* III Kant demonstrates a conceptual connection between the moral law as a principle of reason and free will. He then argues from the possession of theoretical reason to the existence of free will, and from there to the authority of the moral law. (See Hill 1992.) In the *Critique of Practical Reason* Kant relies on the same conceptual connection between morality and free will, but *reverses* the order of argument, moving *from* the authority of the moral law as a "fact of reason" *to* the reality of free will. I sketch the latter argument.

It is part of our ordinary concept of duty or moral requirement that if an action is your duty, you ought to do it regardless of your desire-based interests. We take duties to apply with special normative necessity: they give us reasons for action that do not depend on desires and that limit the force of desire-based reasons that are inconsistent with duty. In this way, moral requirements are "unconditional practical laws" that are reason-giving in virtue of their "legislative form" (CPrR, 5.27). An agent who acts from duty is motivated by the legislative form of the maxim: her reason for acting is the fact that the action is morally required. This analysis leads to the thesis that "freedom and unconditional practical law reciprocally imply each other," which consists of two claims: (1) If an agent is subject to the moral law, that agent is transcendently free; and (2) If an agent is transcendently free, that agent is subject to the moral law (CPrR, 5.28–9). The first is needed to establish free will, and the argument for this claim is simple. An agent subject to the moral law can act from the moral law – that is, can set aside desire-based reasons and take the legislative form of a maxim (the fact that an action is required) as a sufficient reason for action. Desire-based reasons can ultimately be traced back to empirical conditions that produce the relevant desires and interests. Thus, an agent who can base her reasons for action simply on the legislative form of her maxim can act independently of determination by empirical conditions. An agent who can act on this kind of reason satisfies the definition of transcendental freedom.

Are we such agents? Kant claims that the moral law is given as a "fact of reason" (CPrR, 5.31). In ordinary practical reasoning we acknowledge the authority of moral requirements and are conscious of our capacity to act from the moral law. Its authority is reflected in the standards to which we hold ourselves and in the workings of conscience and the moral emotions. By establishing that we are subject to and can act from moral law, the fact of reason establishes that we are free. Indeed, our recognition of the authority of the moral law is at the same time an awareness of freedom: "this fact is . . . indeed identical with consciousness of freedom of the will" (CPrR, 5.42).

Two comments on this way of establishing free will. First, in addition to showing that reason is a causal power in us, Kant's moral theory specifies the normative principle that governs that power. Moral consciousness gives us a determinate positive conception of our free agency: it is the capacity to act from the moral law (cf. *MM*, 6.213–14, 6.224–5). Second, to defend moral responsibility as he understands it, Kant must hold that we act freely even when we do not act from moral reasons. His view is that we have this capacity even when we fail to exercise it. Thus he can say that an action is freely chosen if performed by an agent with the capacity to act from the moral law, whether or not that agent acts from moral reasons.

In order to respect the limits on knowledge established by the first *Critique*, Kant stresses that the reality of free will is not an object of theoretical knowledge, but has "objective though only practical reality" (*CPrR*, 5.48, 5.49). Kant has not given a theoretical proof that we are free, nor is free will given in intuition. Rather the warrant for ascribing free will is our consciousness of the authority of the moral law. The assertion that we have free will is rationally based, but not in the way that knowledge claims are. Because the reality of free will is not established theoretically, it does not expand theoretical knowledge and, for example, it cannot enter into empirical explanations of events. Finally, the capacity with which free agency is identified is specified through moral consciousness – as the capacity to act from the moral law. In sum, both the grounds for ascribing free will to ourselves and our determinate understanding of what free will is are given by moral consciousness, and this idea can only be used from the practical perspective, as part of our self-conception as agents.

References and Further Reading

- Allison, Henry (1990). *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Hill, Thomas, E., Jr. (1992). Kant's argument for the rationality of moral conduct. In *Dignity and Practical reason*. Ithaca, NY: Cornell University Press.
- Kant, Immanuel (1996). *The Metaphysics of Morals, Groundwork of the Metaphysics of Morals, and Critique of Practical Reason*. In *Practical Philosophy*, ed. and tr. Mary J. Gregor. Cambridge: Cambridge University Press.
- Kant, Immanuel (1998). *Religion Within the Boundaries of Mere Reason*. In *Religion and Rational Theology*, eds. and tr. Allen W. Wood and George Di Giovanni. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. (1996). Morality as freedom. In *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Reath, Andrews (1997). Introduction to *Critique of Practical Reason*. In Immanuel Kant, *Critique of Practical Reason*, ed. and tr. Mary J. Gregor. Cambridge: Cambridge University Press.
- Rawls, John (2000). *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press.
- Wood, Allen, W., ed. (1984). *Self and Nature in Kant's Philosophy*. Ithaca, NY: Cornell University Press.