

II. Intelligible character

The usual explanation of Kant's rejection of compatibilism and consequent appeal to an inaccessible noumenal domain or intelligible character is his concern to ground morality. Starting with this assumption, which is itself a reasonable one, if one reads the first *Critique* account of freedom in light of that of the second, it is sometimes also maintained that the main reason for Kant's dissatisfaction with a compatibilist account of freedom lies in his manifestly inadequate view of nonmoral motivation. According to this line of criticism, which can be traced back at least to T. H. Green and has been reaffirmed recently by both Terence Irwin and Allen Wood, Kant believed that all nonmorally motivated actions must be motivated and, therefore, "caused" by the desire for pleasure.¹⁴ Given this belief, Kant could account for the possibility of nonhedonistic or nonegoistic action, as required by his moral theory, only by locating such action in a distinct noumenal world and assigning to it a separate intelligible cause. As Wood succinctly puts it, "the free will problem arises for Kant [in the radical form in which it does] because he is a thoroughgoing psychological hedonist about all the natural causes that might act on our will."¹⁵ The clear implication of this, which is drawn by both Irwin and Wood, is that were it not for this unfortunate thesis in empirical psychology, the way would have been open for Kant to have treated the free will problem along standard compatibilist lines.

I shall argue later (Chapter 5) that Kant does not, in fact, hold such a theory of motivation or, more simply, that he is not a psychological hedonist. But even if he were, it could not account for his analysis of freedom in the *Critique of Pure Reason*. What is distinctive about this analysis is Kant's contention that *both* moral *and* pragmatic or prudential imperatives indicate a causality of reason. Given this, it is clearly a mistake to claim that, in the first *Critique* at least, Kant introduces his noumenalism merely to account for the possibility of acting on the basis of the categorical imperative (as contrasted with the capacity to act on the basis of imperatives in general).

Two passages, one from the Dialectic and one from the Canon, should suffice to make this clear. In the first, shortly after stating that reason imposes imperatives in "all practical matters [*in allem Praktischen*]," which presumably includes prudential as well as moral matters, he contends that "the 'ought' pronounced by reason" is at work "whether what is willed be an object of mere sensibility (the pleasant) or pure reason (the good)." In either case, Kant suggests:

Reason does not follow the order of things as they present themselves in appearance, but forms for itself with perfect spontaneity an order of its own according to Ideas according to which it declares actions to be necessary, even though they have never taken place, and perhaps never will take place (A548/B576)

In the second, after distinguishing between moral laws (which dictate our duty) and pragmatic laws (which dictate the means necessary to the ends

FREEDOM AND RATIONAL AGENCY

stemming from our sensuous nature), he notes that both count equally as “objective laws of freedom which tell us what *ought to happen*, although perhaps it never does happen” (A802/B830).

These passages make it clear that in the *Critique of Pure Reason*, if not in his later works in moral philosophy, Kant regards the capacity to act on the basis of imperatives in general (not merely the categorical imperative) as the defining characteristic of free agency. They also suggest that the spontaneity presumably required to act on the basis of an ought (whether moral or prudential) is the source of Kant’s dissatisfaction with the compatibilist account of agency in terms of empirical character and, therefore, the reason for his introduction of the conception of an intelligible character.

What is not clear at this point is why Kant should think that this introduction is necessary, particularly when one considers its implications. Let us grant, as seems reasonable, that to engage in practical reasoning is to deliberate about what one ought to do (whether in a moral or a prudential sense). Let us further grant that this deliberation requires the “spontaneity” of reason in that it involves the forming of ends or “ought-to-bes” and rules that are not based solely on what one in fact desires at a given moment but rather reflect what one would choose if one were perfectly rational. The problem is that none of this appears to require the abandonment of the compatibilist conception of agency, much less the appeal to any noumenal or “merely intelligible” activities. Thus, we are brought back to our original question: What, apart from specifically moral considerations, is the basis for Kant’s dissatisfaction with the compatibilist account of agency available to him on the basis of the “naturalized” construal of the causality of reason?

In order to understand Kant’s seemingly gratuitous insistence on a merely intelligible moment of spontaneity in the conception of rational agency, we must look not to his moral theory or motivational psychology but rather to his views on the spontaneity of the understanding and reason in their epistemic functions. Indeed, Kant himself indicates the connection when, in a famous passage, he states that through mere apperception we are conscious of certain faculties, namely, understanding and reason, “the action of which cannot be ascribed to the receptivity of sensibility” (A547/B575). Since the operative contrast with the receptivity of the senses for Kant is always the spontaneity of the understanding (and reason), the clear implication is that apperception provides us with a consciousness of this spontaneity and that this consciousness is an inseparable component of the conception of ourselves as cognitive subjects.

Kant’s claim that the understanding is spontaneous (A51/B75) can best be understood in terms of his identification of its fundamental activity with judgment (A73/B94). Largely against the empiricists, he argues that the senses provide the mind with the data for thinking objects, but not with the thought or knowledge thereof. The latter, he maintains, requires the active taking up of the data by the mind, its unification in a concept or synthesis, and its reference to an object. All this is the work of judgment, which is simply the spontaneity of the understanding in action.

EMPIRICAL AND INTELLIGIBLE CHARACTER

A helpful way of explicating what Kant means by the spontaneity of the understanding in its judgmental activity (epistemic spontaneity) is to consider judgment as the activity of “taking as” or, more precisely, of taking something as a such and such.¹⁶ Thus, in the simplest case, an indeterminate something = x is taken as an F , in more complex cases, Fx is qualified by further “determinations” or predicates, for example, Fx is G . In still more complex cases, distinct takings (categorical judgments) are combined in a single higher order “taking” (hypothetical and disjunctive judgments). **The key point, however, is that in all cases the activity of “taking as” is constitutive of judgment.**

judgment & 'taking
as'



This same picture of judgment also clarifies the connection between the spontaneity of the understanding and apperception, which is a central theme of the Transcendental Deduction, particularly in the second edition. Eschewing all details, the main point is simply that although we can perfectly well perceive or intuit x 's that are F 's (such a capacity falls within the sphere of sensibility and can be attributed even to animals), we cannot conceive or represent to ourselves an x as F without not only doing it, that is, consciously taking it as such, but without also in some sense “knowing what one is doing.” This peculiar mode of cognitive self-awareness is what Kant terms “apperception.” As such, **it is not another thing that one does when one judges (a kind of second-order knowing that one is knowing); it is rather an inseparable component of the first-order activity itself.**¹⁷

Apperception, so construed, is contrasted with inner sense, which is the medium through which we can attain to an empirical knowledge of our mental states. Since such knowledge is subject to the transcendental conditions of experience, particularly the form of time, it yields an awareness only of the phenomenal self or the self as it appears to itself. With respect to this self there is no spontaneity. By contrast, apperception, as the consciousness of the act of thinking or, more properly, the self-consciousness built into that very act, is just a consciousness of spontaneity. Thus, in the B-Deduction, Kant describes the representation “I think” (which he identifies with apperception) as itself “an act of spontaneity” (B132), and later, in the Refutation of Idealism, he states, “The consciousness of myself in this representation ‘I’ is not an intuition, but a merely *intellectual* representation of the spontaneity of a thinking subject” (B278).¹⁸ To be sure, the point of the latter claim is to deflate the pretensions of the Cartesian (problematic) idealist to gain knowledge of the self simply by reflecting on the *cogito*. Self-knowledge, for Kant, requires intuition and, as he argues in the Refutation, ultimately outer intuition. Nevertheless, the fact remains that Kant’s account of the understanding commits him not only to the doctrine of the spontaneity of thinking but also to the assumption of a consciousness of this spontaneity that is “merely intellectual” or “intelligible” and that is, therefore, distinguished from self-knowledge, which requires intuition and is based on inner experience.¹⁹

Similar considerations apply, *mutatis mutandis*, to the faculty of reason insofar as its exercise is construed as the logical activity of inference drawing

or reasoning. Although Kant usually has syllogistic reasoning in mind, the main point is simply that any genuinely inferential process (whatever its logical form) involves deriving conclusions from premises in such a way that the premises are taken as justifying the conclusion. In other words, the premises must not only be good and sufficient reasons for asserting the conclusion, they must also be regarded as such. Moreover, as before, this taking as is a spontaneous, inherently self-conscious activity of the subject.

Finally, insofar as reason has a “real use” distinct from that of the understanding, Kant sometimes attributes a distinct, higher level of spontaneity to it in virtue of its total independence of sensibility and its conditions.²⁰ Kant gives a good indication of what he means by this when, in a previously cited passage, he states that “reason . . . frames for itself with perfect spontaneity an order of ideas” (A548/B576). Translating this into the epistemological context suggests the doctrine, which Kant sketches in the Appendix to the *Dialectic*, that reason has the capacity to form ideas and to regulate enquiry in accordance with these ideas. Such formation and regulation involve spontaneity because, rather than simply reflecting a pre-given order of nature, reason projects an order of its own “in accordance with ideas,” that is, it generates a set of norms and goals of explanation in terms of which scientific enquiry must proceed.²¹ This “projection” is, of course, likewise an inherently self-conscious activity.

Given this brief glance at Kant’s views on the spontaneity of thought, we are now in a position to determine why Kant insisted on a spontaneity component in his conception of rational agency and why he thought that the inclusion of such a component requires the introduction of the notion of a nonempirical, merely intelligible character. In both cases, the essence of the answer lies in Kant’s implicit assumption that to conceive of oneself (or someone else) as a rational agent is to adopt a model of deliberative rationality in terms of which choice involves both a taking as and a framing or positing. Since these activities, as expressions of spontaneity, are themselves merely intelligible (they can be thought but not experienced), it is necessary to attribute an intelligible character to the acting subject, at least to the extent to which one regards that subject’s reason as practical.

As presented in the *Critique of Pure Reason*, the distinguishing feature of this model is the virtual identification of rational agency with action on the basis of an *ought*.²² This identification is, of course, intimately related to Kant’s views on moral deliberation and obligation, but it does not involve any conflation of the two. On the contrary, as we have already seen, it is intended to cover both moral and prudential deliberation. In both cases, the essential point is that deliberation involves an appeal to some rule of reason (imperative), which specifies what course of action is “right” or “permissible” in a given situation for an agent, who, as affected by sensible inclination, does not always do what reason dictates ought to be done. Moreover, as the latter point indicates, this model is operative both in the context of deliberation, where it characterizes how one takes oneself qua engaged in a deliberative process, and in the context of appraisal or imputation, where it

grounds judgments of praise and blame (including but not limited to the moral variety) on the basis of the actions or “omissions of reason.”²³

The relevance to this model of the notion of taking as, or more precisely, its practical analogue, becomes apparent when we consider how it construes the relationship between rational agents and their inclinations or desires. The key point here is that even in the case of desire-based actions, a rational agent is not regarded as being determined in a quasi-mechanistic fashion by the strongest desire (roughly the Leibniz–Hume model). On the contrary, to the extent to which such actions are taken as genuine expressions of agency and, therefore, as imputable, they are thought to involve an act of spontaneity on the part of the agent, through which the inclination or desire is deemed or taken as an appropriate basis of action. Moreover, much like the conceptual determination of sensible intuition in the epistemic context, this occurs by subsuming the inclination or desire under a practical rule or principle. Thus, even desire-based or, as Kant later termed it, “heteronomous” action involves the self-determination of the subject and, therefore, a “moment” of spontaneity.

Kant hints at this at the very beginning of his discussion of practical freedom in the *Dialectic*, when he remarks that in considering a free act we are constrained to consider its “cause” as “not . . . so determining that it excludes a causality of our will” (A534/B562). Behind this seemingly paradoxical locution is just the thought that the sensible inclination, which from the point of view of the action’s (and the agent’s) empirical character is viewed straightforwardly as cause, is, from the standpoint of this model, seen as of itself insufficient to determine the will. Moreover, this insufficiency is not of the sort that can be made up for by introducing further empirically accessible causal factors. The missing ingredient is the spontaneity of the agent, the act of taking as or self-determination. Since this can be conceived but not experienced, it is once again something merely intelligible.

Further indications of Kant’s view are contained in a set of *Reflexionen* closely related to the discussion of freedom in the *Critique*.²⁴ In one of them Kant states that actions are to a large part “induced [*veranlasst*] but not entirely determined by sensibility; for reason must provide a complement of sufficiency” (*ein complement der zulänglichkeit*); in another he speaks of reason using but not being determined by the natural condition of the subject (R 5611 18: 252; R 5612 18: 253). Similarly, in a course of lectures stemming from 1794 to 1795, Kant is recorded as claiming that not merely sensible stimuli but also “the concurrence of the understanding” [*die concurrenz des Verstandes*] is necessary to determine the will (MK₃ 28: 1015).²⁵ In spite of their dogmatic metaphysical flavor, all of these texts may be taken merely to be affirming the model of deliberative rationality in terms of which we must conceive ourselves insofar as we regard our reason as practical.

The same may be said of Kant’s fullest and most important published account of his conception of practical spontaneity, which is contained in *Religion within the Limits of Reason Alone*. As he there puts it, “freedom of the will [*Willkür*] is of a wholly unique nature in that an incentive can

FREEDOM AND RATIONAL AGENCY

determine the will to an action *only insofar as the individual has incorporated it into his maxim* (has made it into the general rule in accordance with which he will conduct himself . . .)” (Rel 6: 24; 19). Since this claim, which I call the Incorporation Thesis, underlies virtually everything that Kant has to say about rational agency, we shall return to it throughout this study. For the present, it must suffice to note two features of this thesis that bear directly on our current concerns. First, it makes it clear that for Kant an inclination or desire does not *of itself* constitute a reason for acting. It can become one only with reference to a rule or principle of action, which dictates that we ought to pursue the satisfaction of that inclination or desire. Moreover, as should already be clear from the previous discussion, the adoption of such a rule cannot itself be regarded as the causal consequence of the desire or, more properly, of being in a state of desire. On the contrary, it must be conceived as an act of spontaneity on the part of the agent.

Second, the Incorporation Thesis also enables us to see more clearly the connection, to which Kant himself alludes, between his conception of rational agency and his doctrine of transcendental apperception.²⁶ Thus, in light of this thesis, one may say that just as it must be possible for “I think” to accompany all my representations in order for them to be “mine,” that is, in order for me to be able to represent anything through them, so too it must be possible for the “I take” to accompany all my inclinations if they are to be “mine” qua rational agent, that is, if they are to provide motives or reasons for acting. Again, just as sensible intuitions are related to an object only by being subsumed under concepts, so too sensible inclinations are related to an object of the will only insofar as they are “incorporated into a maxim,” that is, subsumed under a rule of action. Finally, and most significantly for the understanding of Kant’s conception of intelligible character, the “I take,” like the “I think,” can be conceived but not experienced. In other words, I can no more observe myself deciding than I can observe myself judging, although in both cases I must be conscious of what I am doing. That is precisely why both activities are merely intelligible in the specifically Kantian sense.

The relevance of the “positing” or “projecting” function of practical reason is likewise apparent. Just as in the theoretical realm the proper, regulative function of reason is to guide enquiry by framing an ideal order involving the systematic connection of phenomena under laws, so too, in the practical realm, its proper function is to guide conduct by framing an order of ends or ought-to-bes. Like its theoretical analogue, this activity is an expression of the spontaneity of reason because it goes beyond what is dictated by the sensible data, which in this case are the desires and inclinations of the agent. Insofar as one *x*’s because one judges that one ought to *x* (whether for moral or prudential reasons), one *x*’s on the basis of rational considerations. The “necessity” to *x* is, therefore, a rational necessity stemming from “objective laws of reason” (or at least putative laws), not a causal necessity stemming from antecedent conditions. As such, it is compatible with genuine freedom, since it does not follow from the fact that *x*-ing is rationally or objectively

necessary that one will x , whereas it does follow from the fact that it is causally necessary.

In short, both aspects of practical spontaneity are essential to the conception of ourselves as rational agents. I cannot conceive of myself as such an agent without regarding myself as pursuing ends that I frame for myself and that I regard as rational to pursue. Correlatively, I cannot conceive of myself as such an agent without assuming that I have a certain control over my inclinations, that I am capable of deciding which of them are to be acted upon (and how) and which resisted. These are, as it were, necessary presuppositions for all who regard their reason as practical. Kant indicates this in the *Groundwork* by suggesting that we cannot act except under the idea of freedom (Gr 4: 448; 115). In the more metaphysical language of the *Critique of Pure Reason*, this means that we cannot conceive of ourselves as rational agents without attributing to our agency an "intelligible character," capable of determining itself to act on the basis of rational principles, "independently of the conditions of time."

Clearly, it does not follow from the fact (assuming it is a fact) that we are rationally constrained to attribute such a character to ourselves insofar as we regard ourselves as rational agents that we really do possess it. It does follow, however, that we cannot both deny such a character and affirm our status as rational agents. As we shall see in the next chapter, that is precisely why Kant frequently indicates that it remains an open question whether reason has causality.

III. The compatibility of empirical and intelligible character

So far we have considered Kant's conceptions of the empirical and the intelligible character of rational agency separately. We have now to deal with the question of their connection, that is, with the attribution of both characters to a single agent or action. Although the entire account is highly obscure, two passages in particular suggest the air of paradox that Pistorius and so many others have found in Kant's attempt to reconcile causal determinism at the phenomenal level with his incompatibilist conception of freedom.

In the first, after making the previously noted claim that "if we could exhaustively investigate all the appearances of men's wills, there would not be found a single human action which we could not predict with certainty, and recognize as proceeding necessarily from its antecedent conditions," Kant goes on to insist that this does not conflict with the possibility of considering these same actions "in their relation to reason" (A550/B578). Moreover, so considered, they are regarded as actions that ought or ought not to have occurred and, therefore, as free in an incompatibilist sense.²⁷

The second is Kant's notorious discussion of a malicious lie, which he describes as an instance of a voluntary action (*willkürliche Handlung*) and treats as an illustration of the "regulative principle of reason" that it is nec-