

IS KANT'S TRANSCENDENTAL PHILOSOPHY INCONSISTENT?

History of Philosophy Quarterly 8, 1991, pp. 357-372.

Derk Pereboom

Penultimate draft

In the Transcendental Deduction, the centerpiece of his transcendental philosophy, Kant argues that we can come to know that the self is the agent of synthesis and the source of the categories.¹ This position apparently conflicts with another Kantian claim, that nondegenerate (i.e. non-analytic) knowledge of this self is impossible. Since this self is not the empirical self, but the self as it is in itself, Kant is committed to disavowing knowledge of it in virtue of his general denial of knowledge of noumena. This logical difficulty is all the more puzzling in light of Kant's critique of rational psychology in the Paralogisms. There, Kant argues that the rationalist is deluded in his aspirations to knowledge of the self that

1. In this paper Kant's Critique of Pure Reason is cited in the traditional way: 'A' indicates the first edition, 'B' indicates the second edition. All citations are taken, with some alterations, from Immanuel Kant's Critique of Pure Reason, tr. Norman Kemp Smith, (London: Macmillan, 1929). The following is a list of other texts cited in this paper:

- Ak Immanuel Kant, Kant's gesammelte Schriften, edited by the Königliche Akademie der Wissenschaften and its successors, (Berlin: George Reimer (subsequently W. de Gruyter), 1902-).
- Ak VIII Quotations from Ak VIII are from the translation of On a Discovery in The Kant Eberhard Controversy by Henry Allison, (Baltimore: Johns Hopkins, 1973).
- AG G. W. Leibniz, Philosophical Essays, tr. Roger Ariew and Daniel Garber, (Indianapolis: Hackett, 1989)
- AT Oeuvres de Descartes, edited by Ch. Adam and P. Tannery (revised edition, Paris: Vrin/C.N.R.S., 1964-76). Quotations are from The Philosophical Writings of Descartes, translated by John Cottingham, Robert Stoothof, and Dugald Murdoch, (Cambridge: Cambridge University Press, 1985)
- G G. W. Leibniz, Die Philosophischen Schriften, G. J. Gerhardt, ed. (Hildesheim: Georg Olms, 1965).

lies beyond appearance. Yet it would appear that Kant claims transcendental knowledge of this very self in his own positive metaphysical program.

This apparent contradiction at the core of Kant's system has not gone unnoticed by commentators. Among these is P. F. Strawson, who argues that recognition of this very problem should lead us to a radical revision of Kant's picture of philosophy. Strawson puts the problem this way:

A question naturally arises regarding the internal consistency of the 'critical' philosophy and in particular of the doctrines of 'transcendental idealism.' If the spatio-temporality and law-likeness of nature and all that is in it are to be attributed to a subjective source in the human mind, this source cannot be identical with the human mind as a part of nature, the topic of empirical psychology -- a discipline which he rather dismissively describes as a kind of 'physiology of inner sense.' So when Kant speaks of our sensibility and our understanding as the source of space, time, and the categories, it must be to ourselves as we (supersensibly) are in ourselves that he is referring. But this conclusion appears to be in direct contradiction to the doctrine he is at pains to emphasize in his criticism of 'rational psychology', which seeks to argue from the bare fact of self-consciousness to the conclusion that the soul is a simple, immaterial, indestructible substance. The fallacies of 'rational psychology' Kant declares to be among those to which pure reason is prone in its search for the unconditioned which cannot be found in experience. It is not clear why Kant's own brand of what might be called 'transcendental psychology' which is clearly distinguished from the empirical -- does not fall under the same interdict that he places on rational psychology... These

and other difficulties may tempt us to reconstrue, or reconstruct, the central doctrines of the critical philosophy.²

I believe that although Strawson correctly describes the purview of transcendental philosophy to include features of the noumenal self, Kant's attack against rationalism is not compromised. The interpretation I shall advocate has the following structure. A central aim of the Critique is to diagnose the source of rationalist metaphysics' failure as a science, and to replace this pseudo-science with a new metaphysics, a true *scientia*. One measure of the success of this project, in Kant's own terms, is that the new metaphysics--transcendental philosophy--be exempt from the problems of the old. According to Strawson, a main difficulty Kant finds in rationalist metaphysics is that it is about things in themselves, and hence he claims that Kant's attempt to set up a new metaphysics is a failure, since it is still about things in themselves. But this view is mistaken. According to Kant, the deepest problem for rationalist metaphysics is not its claim to knowledge of noumena, but its unchastened apriorism, and correcting this problem does not yield a denial of all noumenal knowledge.

Kant does believe that the domain of transcendental philosophy is limited, yet it is not limited to phenomena. Philosophy is, in part, a realist science: it is a science partly about the self as it is independent of the way we perceive and process information about it--the noumenal self. It is, indeed, important for Kant that a principled distinction can be

2. This quotation is from Strawson's article "Kant's Philosophy of Mind," in The Oxford Companion to the Mind, ed. Richard L. Gregory, (Oxford: Oxford University Press, 1987), p. 408. Strawson explores this line in greater detail in his The Bounds of Sense, (London: Methuen and Co., 1966). See also W. H. Walsh, Kant's Criticism of Metaphysics, (Edinburgh, Edinburgh University Press, 1975), p. 251, and Ralph C. S. Walker, Kant, (London, Routledge and Kegan Paul, 1978), pp. 133-5.

made between the aspects of noumenal self that can be known and those which cannot. Yet the fundamental reason Kant denies knowledge of certain aspects of the noumenal self is that all attempts to justify such knowledge are beset by an insurmountable difficulty. The problem is that all such attempts involve strict a priori justifications, justifications which are completely independent of experience, and we have very good reasons to believe that they cannot produce nondegenerate knowledge in us. Justifications for beliefs about certain other aspects of the noumenal self, by Kant's own transcendental philosophy, do not suffer from this problem, and hence such knowledge does not fall to his critique of rationalist metaphysics.

II

Let us begin by exploring Kant's claim that substantive knowledge of the self as a thing in itself is impossible. As Strawson argues, and I believe correctly, the noumenal self functions as the subject of judgment and as the structuring agent of experience.³ Since for Kant judging and structuring are activities, he calls the noumenal self the determining subject: "In all judgments I am the determining subject of that relation which constitutes the judgment." (B407) But Kant argues for these two epistemological theses about the determining subject: first, that we humans have no intuition of it, i.e., we have no

3. In my view, the self as it is in itself is an aspect of the self, rather than one of two (or three) selves. This interpretation fits with Kant's preferred mode of thinking of things in themselves:

Now let us suppose that the distinction, which our Critique has shown to be necessary, between things as objects of experience and those same things as things in themselves, had not been made. (Bxxvii, emphasis mine)

immediate representation of the determining subject as a particular object, and second, that thus we can have no knowledge (Erkenntnis) of the determining subject as an object.

So first, I have no intuition of myself as determining subject. Rather, I can only intuit myself as an appearance:

...I do not have another self-intuition which gives the determining in me (I am conscious only of the spontaneity of it) prior to the act of determination, as time does in the case of the determinable, I cannot determine my existence as that of a self-active being; all that I can do is to represent to myself the spontaneity of thought, that is, of the determination; and my existence is still only determinable sensibly, that is, as the existence of an appearance. (B158n. cf. B152ff, B407)

Although inner sense or introspection allows me to intuit my mental states, which are the products of activity or determination, it does not reveal the active or determining subject that possess them. Kant follows Hume on this view, (Treatise, p. 207, 633-6), although Kant does not conclude with Hume that there is no such thing as an active subject separate from mental states. My inability to intuit myself as subject of thought leads Kant to claim that neither can I have knowledge (Erkenntnis) of myself as subject. I can only know myself as appearance (i.e. my empirical self):

...in the transcendental synthesis of the manifold of representations in general, and therefore in the synthetic original unity of apperception, I am conscious of myself, not as I appear to myself, nor as I am in myself, but only that I am. This representation is a thought not an intuition. Now in order to know ourselves, there is required in addition to the act of thought, which brings the manifold of every possible intuition to the unity of apperception, a determinate mode of intuition,

whereby this manifold is given... Accordingly I have no knowledge of myself as I am but merely as I appear to myself. The consciousness of self is thus very far from being a knowledge of the self... (B157-8)

If I had an intuition of myself as determining subject, I would be in a position to know the nature of this aspect of the self. But since I have no such intuition, I can only know myself as I appear to myself.

Nevertheless, Strawson is right to maintain that Kant does claim some knowledge of the noumenal self by means of transcendental philosophy. By virtue of the Transcendental Deduction, for instance, Kant claims that we can know that the self as it is in itself is the agent of synthesis, that it organizes experience by means of the categories. For a specific example, in the second edition Deduction Kant argues: "Only so far as I can unite a manifold of given representations into one consciousness, is it possible for me to represent to myself the identity of consciousness in these representations." (B133) The structuring self Kant refers to here could not be the empirical self, since the empirical self is constructed by the functions described in transcendental philosophy, and Kant certainly does not hold that the empirical self constructs itself. Thus this structuring self must be the self as it is in itself. Furthermore, in this passage, and throughout the Deduction, Kant is claiming genuine and full-fledged knowledge of this self. The Transcendental Deduction is the most important element of the revolutionized metaphysics which Kant claims to be a true scientia (Bxv-xvii), and it would be absurd to claim that important elements of this true scientia do not count as genuine knowledge.

Some interpreters answer the objection Strawson raises by having Kant admit only phenomenal or empirical self-knowledge, while downplaying his apparent endorsement of

noumenal self-knowledge. Karl Ameriks, for example, attempts to solve Strawson's problem (as presented by R. C. S. Walker) by claiming that "we ought not to conflate the transcendental conditions of knowledge with properties of the self in itself."⁴ He maintains that in Kant's view, although we can say that "the structure of human knowledge depends on 'our' categories and forms of intuition," we cannot say that it depends "on the existence or action of any known individual human selves." Ameriks is clearly right insofar as his claim is that in Kant's theory, the structure of human knowledge does not depend on a known empirical self. But Ameriks also says that "transcendental structures... are not asserted to be due to a literal and personal action."⁵ Kant's frequent use of the term 'I' to refer to the subject of synthesis counts as positive evidence against Ameriks' view, and I believe that there is no convincing textual evidence that the subject of synthesis is not an individual human self.

As Ameriks admits, a version of Strawson's problem arises for his own interpretation since he admits that belief in transcendental structures "may appear to involve instances of knowledge that are contrary to the general prohibitions of Kant's transcendental idealism."⁶ Ameriks suggests that belief in transcendental structures can be legitimized in the same way as some of the speculations of dogmatic metaphysics are.⁷ In Kant's view, such speculations are intelligible, useful, often rest on non-theoretical,

⁴ Karl Ameriks, Kant's Theory of Mind, (Oxford, Oxford University Press, 1982), pp. 287-8, cf. p. 277; Walker, pp. 133-5.

⁵ Ameriks, p. 289.

⁶ Ameriks, p. 289.

⁷ Ameriks, p. 290.

practical premises, but do not amount to knowledge. According to Kant, however, belief in transcendental structures is not speculative metaphysics. The claims about transcendental structures argued for in the Transcendental Deduction, for instance, are not merely intelligible and useful, while not counting as knowledge, and they do not rest on non-theoretical premises. Rather, such claims are instances of transcendental philosophy, which Kant, by his Copernican revolution, aims to justify as genuine knowledge.

In addition to transcendental self-knowledge, Kant allows some quasi-knowledge of the self as it is in itself by immediate consciousness or through cogito-style arguments. As we have seen, Kant says that in the synthetic original unity of apperception, I am "conscious (bewußt) of myself, not as I appear to myself, nor as I am in myself, but only that I am." (B157, cf. B422-3), although he denies that such consciousness has the status of full-fledged knowledge. Kant seems to find the reconciliation of this type of cognition with his critique of rationalism a more pressing problem than the accommodation of the self-knowledge gained through transcendental philosophy.

Thus, although Kant denies knowledge of the noumenal self as determining subject, he also argues that we can know that this self is the structuring agent of experience, and in addition, he maintains that we have some type of cognition of this self through immediate consciousness. Thus the problem that Strawson raises is a real problem for Kant, and demands a solution.

III

To dissolve our apparent inconsistency, I shall argue first, that the domain of the type of noumenal self-knowledge Kant denies is distinct from the domain of the kind he

allows, and second, that admitting this self-knowledge does not require rationalist procedures for justification, and hence does not undermine his critique of rationalism.

Let us first focus on the question of domain. By rejecting the view that we have intuition of the determining subject, Kant is in effect denying that we have immediate cognition of it as an object, for an intuition is by definition immediate knowledge of an object. Moreover, Kant explicitly forswears any knowledge of the determining subject as an object, whether inferred or immediate. This puzzling and obscure denial is a prominent theme throughout the Paralogisms. For example,

Modi of self-consciousness in thought are not by themselves concepts of objects (categories), but are mere functions which do not give thought an object to be known, and accordingly do not give even myself as object. (B406-7, cf. A346=B404, A350, A354, A356, A397, A399, A402, B408-9)

Furthermore, transcendental knowledge is not knowledge of the self as an object:

I entitle transcendental all knowledge which is occupied not so much with objects as with the mode of our knowledge of objects in so far as this mode of knowledge is to be possible a priori. A system of such concepts might be entitled transcendental philosophy. (All-12=B25)

From these passages we might conjecture that the kind of self-knowledge Kant denies is knowledge of the noumenal self as an object, and the kind he allows is not knowledge of this self as an object. But what does this distinction amount to?

Kant, regrettably, does not clearly define the notion of knowledge of the self as an object. Yet, in the Paralogisms, he does distinguish several specific kinds of self-knowledge that would count as knowledge of this self as object. And, as I shall now argue, all of these specific kinds of self-knowledge amount to knowledge of the features of the self that are

characteristic of a classical rationalist substance. Let us call such knowledge rationalist substance-knowledge.

Kant's aim in the First Paralogism is to undermine the rationalist demonstration that the self is a substance. He allows that we can know that the self is a substance in the weak sense of a subject of judgments, but not in the strong sense that "I, as a thinking being, persist for myself, and do not in any natural manner arise or perish." (A352)⁸ In this strong sense, a substance is characterized primarily by independence; a substance is something which is independent of natural causes for its coming to be, its preservation, and its going out of existence. This strong sense of 'substance' has roots in Descartes, who first defines substance as "a thing which exists in such a way as to depend on no other thing for its existence." Seeing that God alone would be a substance on this definition, Descartes liberalizes his notion of substance to include things that are dependent only on God for their existence: "In the case of all other substances, we perceive that they can exist only with God's concurrence." (Principles I, 51, AT VIIIA, 24) Leibniz, in the Conversation of Philarète and Ariste, revises this definition to eliminate abstract things such as properties. For him, a substance is a concrete thing whose existence is independent of any concrete created thing outside of it. (AG 259-63=G VI 581-7) Kant captures the essence of this notion when he equates a substance with a self-subsistent being. (B407)

Further, Leibniz maintains that activity is an essential feature of substance, that is, substances are fundamental sources of causal power. (AG 40-1=G432-3; AG 139-40=G IV 478-9) This feature of substances is linked to the root notion of independence; part of what

⁸ Descartes' Geometrical Exposition is a source for this sense of 'substance.', (AT VII, 161)

it means for a thing to be independent is for it to have its own causal power. Although the causal power of created things ultimately derives from God, nevertheless, within the created order their causal power is basic. We might therefore characterize a classical rationalist substance as a fundamental cause--such a substance depends on no created thing for its existence, while it is a source of causal power in the created world.

Furthermore, Kant claims that the argument of the First Paralogism, which attempts to demonstrate a priori that the noumenal self is a substance, reveals the most fundamental rationalist contention about the nature of the self; the "whole procedure of rational psychology is determined by [this] paralogism." (B411) The reason behind this claim is that the features the other Paralogisms attribute to the self are also characteristics of a rationalist, and in particular, a Leibnizian substance. Let us consider each of the other Paralogisms to determine why Kant's contention is plausible.

In the argument of the Second Paralogism, the rationalist attempts to establish that the self is simple. Leibniz argues that substances must be simple because compound things are dependent on their parts for their existence. (AG 142=G IV 482-3; AG 262=G VI 585-6) If the self were not simple, it would lack the sort of independence necessary for being a substance. The argument of the third Paralogism aims to establish the permanence of substances. Leibnizian substances possess this characteristic because they cannot naturally go out of existence. If they could naturally go out of existence, then they would be dependent on the natural order for their existence, and thus would lack the kind of independence required for being a substance. The argument of the Fourth Paralogism (in the second edition) aims to show that thinking beings "are conscious of their existence as separate and distinct from all matter." (B409) Leibnizian substances are immaterial

because material things are dependent on their parts for their existence, and again, if they were so dependent, they would lack the independence essential to substances. We can thus conclude that in the Paralogisms, where Kant denies knowledge of I as object, all of his limitations on self-knowledge add up to a denial of rationalist substance-knowledge of the self.

This account of the type of self-knowledge Kant denies can be deepened by reflecting on his argument against Leibnizian rationalism in the Amphiboly of Concepts of Reflection.⁹ One of Leibniz's characteristic doctrines is that all substances have natures that consist only in intrinsic and not in extrinsic properties.¹⁰ (By definition, intrinsic properties are non-relational, extrinsic properties are relational.¹¹ For example, being a thinker is intrinsic to Kant, whereas being southwest of Riga is an extrinsic property of Königsberg.) In the Amphiboly, Kant argues that because we are only human, we cannot have insight into the intrinsic natures of things by means of purely intellectual cognitions; because we are human, we cannot have strict a priori knowledge of the natures of Leibnizian substances:

9. For a more thorough exposition of the Amphiboly, see my "Kant's Amphiboly," in Archiv für Geschichte der Philosophie, 1991.

10 Kemp-Smith translates 'das Innere' as 'the inner,' 'innerlich' as 'inner' or as 'inward,' and 'inneres' as 'internal.' In the interests of uniformity, and since the term 'intrinsic' is the current equivalent in the discussion of Leibniz, I have revised his translation accordingly. Similarly, Kemp-Smith renders 'das Äußere' as 'the outer' and 'äußere' as 'outer;' I have substituted 'the extrinsic' and 'extrinsic.'

11. See David Lewis, "Extrinsic Properties," Philosophical Studies 44 (1983), pp. 197-200, for an attempt at a modal definition of these notions. One problem with his view is that if a person is essentially the child of her parents, then her existence will turn out to be an extrinsic property, which is counterintuitive. As far as I can see, the only alternative to defining these notions modally is to regard them as basic.

If by the complaints--that we have no insight into the intrinsic [nature] of things--it be meant that we cannot conceive by the pure understanding what the things which appear to us may be in themselves, they are entirely illegitimate and unreasonable. For what is demanded is that we should be able to know things, and therefore to intuit them, without senses, and therefore we should have a faculty of knowledge altogether different from the human, and this not only in degree but as regards intuition likewise in kind---in other words, that we should not be men but beings of whom we are unable to say whether they are even possible, much less how they are constituted. (A277-8=B333-4)

It is particularly striking that in this passage Kant identifies purely intellectual understanding of things in themselves with insight into the intrinsic nature of things to the noumenal self. Thus, at least sometimes, by 'knowledge of the self as it is in itself' Kant means to refer to knowledge of the self's intrinsic nature, of the non-relational properties that underlie the relational properties we perceive. And again, knowledge of these intrinsic properties, which Kant rejects, would be rationalist substance-knowledge, i.e., knowledge of features essential to substances, in the rationalist conception.

It is also significant that in the subsequent passage Kant links this notion of an intrinsic nature to that of a fundamental causal explanation. The most basic causal explanation for natural phenomena is to be found in the mind, but we cannot know the intrinsic nature of the mind that would yield such an explanation:

even if the whole of nature were revealed to us, we should still never be able to answer those transcendental questions which go beyond nature. The reason of this is that it is not given to us to observe our mind (Gemüt) with any other intuition

than that of inner sense; and that it is precisely in the mind that the secret source of sensibility is located. The relation of sensibility to an object and what the transcendental ground of this [objective] unity may be, are matters undoubtedly so deeply concealed that we, who after all know even ourselves only through inner sense and therefore as appearance, can never be justified in treating sensibility as being a suitable instrument of investigation for discovering anything save still other appearances--eager as we yet are to explore their non-sensible cause. (A278=B334, emphasis mine)

Kant agrees with Leibniz that the intrinsic nature would provide the most basic causal explanation for various features of a substance, such as its capacities and activity. Intrinsic natures, therefore, are fundamental causes, but we can have no knowledge of them.

Thus the import of these passages from the Amphiboly is that we are excluded from knowing the intrinsic natures of things, the natures of Leibnizian substances, which are fundamental causes in the world. The reason why we cannot possess such knowledge is that it would have to be purely intellectual, completely independent of experience for its justification. Kant argues that human beings cannot possess knowledge of this kind.

Support for the view that Kant is most concerned to reject rationalist substance-knowledge is found in his further challenges to rationalism. To the extent that the other anti-rationalist arguments in the Dialectic aim at restricting knowledge of things in themselves, they also seek to undermine claims to knowledge of fundamental causes. In the Antinomy, Kant argues that we can have no knowledge of the intrinsic causal nature of matter, of the ultimate springs and principles of action, or of a necessary being who is the

cause of the world.¹² In the Ideal, Kant attempts to show that we cannot come to know God, the alleged ultimate cause of all things. Thus, there is good reason to believe that his main problem with the domain of all of these arguments is they are supposed to yield knowledge of fundamental causes.¹³

If this account is correct, then the domain of transcendental knowledge is exempt from Kant's attack against rationalism in the Dialectic. By transcendental philosophy we do not acquire rationalist substance-knowledge of the self, that is, we do not come to know the intrinsic nature, and hence the fundamental causal powers, of the self. Rather, transcendental philosophy provides us with knowledge of some of the capacities of the I, and of the use of these capacities in the structuring of experience.

One objection to this account is that these capacities themselves--the forms of intuition and the pure concepts of the understanding--are fundamental causal powers of the self. These *Vorstellungen* do not seem to differ in kind from Leibnizian innate ideas, and according to Leibniz innate ideas are part of the self's intrinsic and fundamental causal

12. The first Antinomy, on whether the world has a beginning in time or space, is not about fundamental causes, but according to Kant neither is it about things in themselves. According to the rationalist position Kant is attacking, time and space are things in themselves; but in his own conception they are appearances.

13. There is a significant affinity among the views of Kant, Locke, and Hume on the issue of knowledge of fundamental causes. In *An Essay Concerning Human Understanding*, IV, iii, Locke argues that although we can have knowledge of certain observable properties, and some features of the corpuscular nature of physical objects, we cannot know their fundamental causal structure. In *An Enquiry Concerning Human Understanding*, Section IV, Part I, Hume states that "elasticity, gravity, cohesion of parts, communication of motion by impulse" are "probably the ultimate causes and principles which we shall ever discover in nature," but the "ultimate springs and principles are totally shut up from human curiosity and enquiry." Thus, Kant's skepticism about fundamental causes is an endorsement of a feature of the modern empiricist tradition.

nature. However, Kant quite clearly attempts to avoid this conclusion. In *On a Discovery*, Kant rejects the view that the a priori intuitions and the categories are causally fundamental. Kant denies that his system allows any innate representations: "the Critique admits of no divinely implanted (*anerschaffene*) or innate (*angeborene*) representations. It regards them all, whether they belong to intuition or to the concepts of the understanding as acquired (*erworben*)." (Ak VIII, 221) He goes on to claim that the forms of intuition and the synthetic unity of the manifold in concepts are acquired in a special way, through an original acquisition:

There is, however, an original acquisition (as the teachers of natural right formulate it), consequently also of that which previously did not exist, and therefore did not pertain to anything before the act. Such is, as the Critique shows, first of all, the form of things in space and time, secondly, the synthetic unity of the manifold in concepts; for neither of these is derived by our faculty of knowledge from the objects given to it as they are in themselves, but rather it brings them out of itself a priori. (Ak VIII, 221-2)

It is crucial that Kant does not profess to know the nature of this original acquisition nor the nature of its source. In particular, he does not provide an account of the fundamental cause of this acquisition. Kant does affirm that such a fundamental cause exists, and that it is innate: "There must, however, be a ground in the subject which makes it possible for these representations to originate in this and no other manner, and which enables them to be related to objects which are not yet given. This ground at least is innate." (Ak VIII, 221-2) But this is not to say that we have specific knowledge of the nature of this fundamental cause, and Kant never affirms that we have such knowledge.

A second objection to my interpretation is that the admission of any knowledge of things in themselves undermines Kant's notion of moral faith. In the Preface to the second edition of the Critique, Kant argues that one advantage of the critical philosophy is that it restricts knowledge of things in themselves to allow us to think we are free in the sense required for moral responsibility. (Bxxx) Thinking that we are free is a condition of viewing ourselves as morally responsible beings. But we cannot (reasonably) think that some knowable aspect of ourselves is free, since science shows that everything of which we have knowledge is causally determined. We can view ourselves as morally responsible beings only if we regard the knowable realm as mere appearance, and if we think of an unknowable aspect of ourselves as free. In the interests of morality, we must think of ourselves as free as we are in ourselves. (Bxxix) But then, if we were to claim knowledge of things as they are in themselves, we would be giving up the room we need to speculate that we are really free.

However, the transcendental knowledge of the self as it is in itself that Kant admits does not undermine moral faith. In Kant's view, freedom is a matter of the fundamental causal origins of human action. If an act is free, then one strand in the causal history of the act ends in the undetermined choice of the agent. The agent is the fundamental cause of a free action because the agent causes the action and is not caused to cause the action. (A533=B561) But since the knowledge of the noumenal self Kant allows does not encompass fundamental causes, moral faith escapes unharmed.

A final problem for this account is that although Kant maintains that the categories (cause and effect, in particular) do not apply beyond the bounds of experience, we want to grant him knowledge of noumenal causal processes. Ameriks suggests that the categories

have a meaning which is independent of their justified use in our spatio-temporal world, and that this is consistent with allowing indirect arguments that there exist things in themselves, and that in some way, which we cannot determine specifically, they are categorized as causal or substantial.¹⁴ I believe that something like this must be right. Furthermore, most important for Kant in his critique of rationalism is first, the denial of the knowledge of features of things in themselves that are characteristic of classical rationalist substances, and second, the rejection of the source of the problem, the type of strict a priori knowledge championed by his rationalist opponents. Some knowledge of causal powers of noumena is consistent with these critical views, and it fails to subvert Kant's case against his rationalist predecessors.

IV

Kant's overarching reason for rejecting knowledge of things in themselves is the failure of the rationalist procedure for justifying such knowledge. I shall argue that in Kant's view, the kind of justification transcendental self-knowledge requires is not rationalist, and hence, allowing this sort of self-knowledge does not undermine the critique of rationalism. This is so because rationalist metaphysics is a priori in the strict sense that it is justified completely independently of experience, whereas the conditions on the justification of transcendental philosophy are not so exacting. Although Kant maintains that transcendental philosophy is a priori, it is only a priori in weak sense, a sense that allows for some appeal to experience.

¹⁴ Ameriks, p. 290; cf. Walker, p. 131.

According to Kant, characteristic of rationalism is the view that we can assume the transcendent perspective, a point of view from which, through reason alone, by overcoming the confusion of sensation, we can apprehend the world, other rational beings, and ourselves, as they are in themselves. At the transcendent standpoint, the process by which we gain knowledge is a priori in the strict sense: it is completely independent of experience for its justification. Yet those who are convinced that we are capable of achieving this standpoint have failed meet their own expectations. First, there is no convergence in the rationalists' specific procedures for acquiring knowledge of things in themselves; "the various participants are unable to agree in any common plan of procedure." (Bvii) Second, no body of successful results exists; "after elaborate preparations, frequently renewed, it is brought to a stop immediately it nears its goal." (Bvii) Kant elaborates on the first point in the Antinomy, by attempting to show that there are proofs, faultless by rationalist criteria, for several sets of opposing conclusions about things in themselves. He expands on the second point by exposing what he thinks to be fatal flaws in fifteen rationalist arguments for conclusions about things in themselves; the four Paralogisms about the self as it is in itself, the eight arguments in the Antinomy concerning the cosmos as it is in itself, and the three arguments in the Ideal for the existence of God. In Kant's analysis, the root of these problems lies in rationalist procedures for justification. The rationalist error is the failure to realize that all genuine knowledge requires an appeal to intuition for its justification.

If Kant's project is to succeed, transcendental philosophy cannot be a priori in the way rationalist metaphysics is. What is Kant's theory of the justification of his new metaphysics?¹⁵ Let us first consider a passage from the Discipline of Pure Reason:

In transcendental knowledge, so long as we are concerned only with concepts of the understanding, our guide is the possibility of experience. Such proof does not show that the given concept (for instance, of that which happens) leads directly to another concept (that of a cause): for such a transition would be a saltus which could not be justified. The proof proceeds by showing that experience itself, and therefore the object of experience, would be impossible without a connection of this kind.

Accordingly, the proof must also at the same time show the possibility of arriving synthetically a priori at some knowledge of things which was not contained in the concepts of them. (A783=B811, last emphasis mine; cf. A61-2=B86, Ak XII 370-1)

Although several commentators have argued that transcendental philosophy is analytic, this passage clearly indicates that it is synthetic.¹⁶ The justification of synthetic knowledge must reach beyond an analysis of concepts. But what, besides conceptual analysis, is involved in such justification? In this passage, Kant also indicates that the way we arrive at transcendental knowledge is a priori. Hence, he must walk a tightrope: justification in transcendental philosophy is a priori in some significant sense, but it must also involve

15. For a more thorough treatment of this question, see my "Kant On the Justification of Transcendental Philosophy," *Synthese* 85, n. 1, October 1990.

16 For example, Ermanno Bencivenga, in *Kant's Copernican Revolution*, (Oxford University Press, Oxford, 1987).

some intuitive element. Since such justification does not plausibly involve a priori intuitions, it would seem that it must depend on experience.

In the above passage, Kant also indicates that the justification of a transcendental proposition proceeds by showing how experience would be impossible if the proposition were not true. Kant's actual procedure in the Transcendental Deduction is consonant with this description. For instance, in the argument from above at A115ff he claims that the self-conscious aspect of our experience would be impossible unless a synthesis by means of a priori concepts takes place, and in the argument from below at A119ff he argues that unless there is such a synthesis, certain features of the objectivity of experience would be impossible.

These transcendental justifications require some appeal to experience. However, the extent to which experience is involved must somehow be limited, for otherwise the apriority of transcendental knowledge is jeopardized, and this is contrary to Kant's intentions. Another passage from the Dialectic indicates that the apriority of transcendental knowledge is preserved by a restriction to possible experience:

...in the Transcendental Logic, although we can never pass immediately beyond the content of the concept which is given us, we are nevertheless able, in relation to a third thing, namely possible experience, to know the law of its connection with other things, and to do so in an a priori manner. (A766=B794, second emphasis mine, cf. A783=A811, quoted above)

One might think that our knowledge of the features of possible experience does not require any appeal to experience at all, because knowledge of possibilities is strictly a priori. But in Kant's view, although merely logical possibilities, propositions or states of affairs that are

not self-contradictory, can be known independently of all experience, real possibilities cannot, and Kant consistently disparages merely logical possibility as irrelevant to the philosophical investigation of reality (B302-3, note added in B; cf. A597=B625, A218=B265ff). Real possibility is important to such investigation, but to know whether a proposition or a concept is really possible an appeal to experience is required. In his discussion of the ontological argument, Kant makes this remark about the possibility of a supreme being:

The analytic criterion of possibility, as consisting in the principle that bare positives (realities) give rise to no contradiction, cannot be denied to it. But the realities are not given to us in their specific characters. Even if they were, we should still not be in a position to pass judgment, since the criterion of the possibility of synthetic knowledge is never to be looked for save in experience... (A602=B630)

We cannot know independently of all experience whether the concept of the supreme being is of a really possible entity, since knowledge of real possibilities requires empirical justification.

Given that the justification of all transcendental knowledge must appeal to really possible experience, and that knowledge of real possibility requires an appeal to experience, in what sense is transcendental justification a priori? Particular experiences are required for the knowledge that digging straight down from New York and hitting the South Pacific is a real possibility, and for the knowledge that it is not really possible that a donkey and a horse produce fertile offspring. So how is transcendental knowledge different from the ordinary empirical variety? Kant's answer is provided in a passage from the Principles:

Synthetic a priori judgments are thus possible when we relate the formal conditions of a priori intuition, the synthesis of imagination and the necessary unity of the synthesis in a transcendental apperception, to a possible empirical knowledge in general. (A158=B197, emphasis mine)

Kant maintains that possible empirical knowledge in general is required to justify the synthetic a priori judgments of transcendental philosophy. 'Possible empirical knowledge in general' plausibly means knowledge which is justified on the basis of any experience possible for us, that is, knowledge based on information derivable from any reasonably long stretch of experience a human being might have. Such justification is a priori because it is independent of any particular experiences, in contrast to knowledge in chemistry and biology, for example, which is dependent on particular experiments and observations. Yet this sort of justification is a priori in a weaker sense than the strict type advocated by the rationalists, since it not completely independent of experience.¹⁷

Support for this view can be found in a passage from the Architectonic of Pure Reason, where Kant attempts to explain how we can acquire transcendental knowledge of certain kinds of synthetic a priori principles about experience:

First of all, how can I expect to have knowledge a priori (and therefore a metaphysics) of objects in so far as they are given to our senses, that is, given in an a posteriori manner? And how is it possible to know the nature of things and to arrive

¹⁷ For a related characterization of a priori knowledge, which inspired the one I present, see Philip Kitcher's "How Kant Almost Wrote 'Two Dogmas of Empiricism'," Philosophical Topics, (vol. 12, no. 2, Fall, 1981, pp. 217-249); Kant's Philosophy of Science," in Self and Knowledge in Kant's Philosophy, (Ithaca: Cornell University Press, 1984), pp. 157-173; "A Priori Knowledge," Philosophical Review 89 (1980): 3-23; and The Nature of Mathematical Knowledge (Oxford: Oxford University Press, 1984).

at a rational physiology according to principles a priori? The answer is this: we take nothing more from experience than is required to give us an object of outer or of inner sense. The object of outer sense we obtain through the mere concept of matter (impenetrable, lifeless extension), the object of inner sense through the concept of a thinking being (in the empirical inner representation, 'I think'). As to the rest, in the whole metaphysical treatment of these objects, we must entirely dispense with all empirical principles which profess to add to these concepts any other more special experience, with a view to our passing further judgments upon the objects. (A847-8=B875-6)

The justification of this type of transcendental knowledge requires empirical information about the objects of inner and outer sense, and the acquisition of this information depends on empirical intuition. However, no particular or special empirical intuitions are needed to acquire this information. On the contrary, it can be derived from any possible human experience. My interpretation of the nature of transcendental justification is also confirmed by an examination of the premises Kant employs in his transcendental arguments, such as 'my experience manifests regularities sufficient for association,' and 'I can become conscious of each of my representations.' It is implausible that such premises are completely non-empirical, although the experience required to gain knowledge of them is so minimal that any possible human experience would suffice.

Hence, Kant's account of justification in transcendental philosophy is consistent with his attack against rationalist metaphysics. Justification in rationalist metaphysics is strictly a priori; no appeal to experience is permitted. In transcendental philosophy, however, justification is empirical, although the appeal made to experience is only minimal.

As I have mentioned, Kant seems more concerned about the status of his claims to cognition of the self by immediate consciousness and by the cogito argument than he is about transcendental knowledge of the self. Perhaps this is because it seems that immediate consciousness of the I and the cogito argument do not appeal to experience, and hence such cognition is apparently acquired by strictly a priori procedures. Kant's struggle with this problem is most evident in the long footnote in the Paralogisms of the second edition at B422-3. Here Kant maintains that the 'I' of 'I think' and 'I exist' refers to the self as it is in itself. But cognition of these propositions demands an appeal to some experience or other:

...when I have called the proposition 'I think' an empirical proposition, I do not mean thereby, that the 'I' in this proposition is an empirical representation. On the contrary, it is purely intellectual, because belonging to thought in general. Without some empirical representation to supply the material for thought, the actus, 'I think,' would not, indeed, take place; but the empirical is only the condition of the application, or of the employment, of the pure intellectual faculty. (B423n, cf. B158-9)

Kant aims to solve the problem by arguing that without some experience or other, plausibly any possible human experience, we could not have access to the proposition 'I think,' although the representation of the I itself is purely intellectual. It is not clear whether Kant would say that the justification of 'I think' has any empirical basis, but perhaps he would allow that such justification requires a minimal appeal to experience.

In summary, we can reply to Strawson that although transcendental philosophy does claim some knowledge of noumena, it is nevertheless consistent with Kant's attack against rationalist metaphysics. Kant's new scientific metaphysics claims no rationalist substance-knowledge of noumena, and more fundamentally, it relies on no strict a priori procedures of justification. This account is consistent with the passage which provides the best support for Strawson's dismissal of Kant's project:

Indeed, it would be a great stumbling-block, or rather would be the one unanswerable objection, to our whole critique, if there were a possibility of proving a priori that all thinking beings are in themselves simple substances, and that consequently (as follows from this same mode of proof) personality is inseparable from them, and that they are conscious of their existence as separate and distinct from all matter. For in this way (*Denn auf diese Art*) we should have taken a step beyond the world of sense, and have entered into the field of noumena; and no one could then deny our right of advancing yet further in this domain, indeed of settling in it, and, should our star prove auspicious, of establishing claims to permanent possession. (B409-410, emphasis mine)

In transcendental philosophy, Kant makes no claims to the type of self-knowledge that he denies to the rationalist, for the noumenal self-knowledge he admits is not rationalist substance-knowledge. Moreover, transcendental philosophy does not lead us beyond the world of sense, since it is minimally dependent on human experience for its justification. Thus, although we do enter the field of noumena, we do not enter it "in this way," that is, by a strict a priori justificatory process. Consequently, Kant can consistently deny our right of advancing yet further in the noumenal domain.

Strawson is right to indicate that the consistency of transcendental philosophy with the polemic against rationalism is of great significance to the project of the Critique of Pure Reason, since here Kant aims to provide foundations for a philosophical science of mind which do not suffer from the problems of rationalist metaphysics. But transcendental philosophy is exempt from these problems because it eschews justifications that do not make any appeal to experience. Yet it is not simply an empirical science. First, it is realistic; it is not just a science about phenomena, as the empirical sciences are, but about selves as they are independent of our perception and mental processing; and second, it is a priori in the weaker sense I have outlined. Hence, in Kant's scheme, philosophy is raised to the level of a true scientia, while its distinctive status, separate from the empirical sciences, is maintained.¹⁸

¹⁸ I wish to thank David Christensen, Hilary Kornblith, Arthur Kuflik, William Mann, and George Sher for helpful comments and discussion.