

The Problem of Self-Knowledge in Kant's "Refutation of Idealism": Two Recent Views

JONATHAN VOGEL
Amherst College

Kant's argumentation in the "Refutation of Idealism" section of *The Critique of Pure Reason* is—to put it mildly—highly compressed. In fact, a step necessary to the completion of the Refutation appears to be missing altogether. If the self can be directly known to persist through change, the Refutation fails, yet Kant seems not to address such a possibility. This apparent omission has recently drawn the attention of scholars, including Henry Allison and Paul Guyer. Allison and Guyer both maintain that Kant had the resources to fill the seeming gap in the Refutation, although they disagree as to exactly what form the completed argument would take. The proposals made by Guyer and Allison do much to illuminate Kant's thinking, but neither, I believe, is fully successful.

A. *The Apparent Gap in the Refutation*

The purpose of the Refutation is to demonstrate the untenability of a kind of skepticism about the external world, which Kant calls "problematic idealism." For the problematic idealist, our knowledge of ourselves is unquestionable, but it is open to doubt whether we have knowledge of anything *except* ourselves. Kant thinks that this skeptical position is ultimately inconsistent, because it turns out that even minimal self-knowledge (consciousness of oneself as "determined in time") presupposes knowledge of the external world (B 274–75).¹

¹ I follow the standard convention for references to the *Critique of Pure Reason*. Translations from the *Critique* are from Norman Kemp Smith's edition (New York: St. Martin's Press, 1974). Other translations used are *Anthropology From A Pragmatic Point of View* trans. Mary Gregor (The Hague: Martinus Nijhoff, 1974); the *Critique of Judgment*, trans. Werner Pluhar (Indianapolis: Hackett Publishing Co., 1987); and the *Nova Dilucidatio (A New Exposition of the First Principles of Metaphysical Knowledge)* trans. John Reuscher, in Lewis White Beck, *Kant's Latin Writings* (New York: Peter Lang Publishing, 1986). Citations are to these translations; I also provide corresponding refer-

In trying to establish this dependence, Kant apparently proceeds as follows (B 275–76):

- (A) I have knowledge of myself: “I am conscious of myself as determined in time.”
- (B) In order to have such knowledge, I must have knowledge that something “permanent” existed from the time of my previous mental state through to the present.
- (C) This permanent could not be one of my mental states.
- (D) Therefore, the permanent of which I have knowledge must be something “outside me,” i.e., an object in the external world.

The exact meaning and the cogency of this argument’s premises are open to dispute, but on almost any account (A) through (C) would not be sufficient to support (D). (B) recapitulates a theme from the First Analogy, according to which knowledge involving “time-determination” requires the ascription of properties to a permanent object or objects.² Hence, in particular, consciousness of oneself as determined in time will require the recognition of some permanent object(s) to which properties are attributed. Now, this object could be external to the mind, but why must it be? Why not suppose that a person’s knowledge is limited to the ascription of properties to him- or herself, with the self serving as the known permanent object necessary for time-determination? Unless this possibility is somehow excluded, it would seem that a problematic idealist could grant Kant’s premises and still not concede the possibility of knowledge concerning objects in the external world.

Thus, to the extent that (A) through (D) captures the content of the passage at B 275–76, there is a critical gap in Kant’s argument. To refute the problematic idealist, Kant would need to supply some additional reasoning to establish that it is impossible to know the self as permanent, if the self is

ences to the *Akademie* edition of Kant’s complete works, *Kants Gesammelte Schriften*, edited by the Königlich-Preussischen Akademie der Wissenschaften (Berlin: Walter de Gruyter, 1929—). The interpretations I discuss are found in Henry Allison, *Kant’s Transcendental Idealism* (New Haven: Yale University Press, 1983) and Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987). These works are cited by page number, as “Allison” and “Guyer.”

² Since my intention is to focus on a particular aspect of the Refutation, I am intentionally unspecific about the exact meaning of such terms as “time-determination.” I use the word “permanent” as synonymous with “persisting,” to mean “continuing to exist throughout an interval of time.” Of course, Kant thinks that the First Analogy establishes that time-determination requires knowledge of objects that are permanent in a much stronger sense than this.

the only object of knowledge. Or, as I will be saying, Kant had the problem of finding a way to complete the Refutation.

B. Allison's Humean Interpretation

Henry Allison has presented an interpretation of the Refutation according to which the argument there would be completed by an appeal to special considerations about inner sense:

Kant's claim is that the required permanent cannot be something inwardly intuited, and this is equivalent to the claim that it cannot be an object of inner sense. This follows from Kant's essentially Humean view of inner intuition or experience and its object. Once again, all that we inwardly intuit is the appearing (to ourselves) of our own representations. There is no additional intuition of a subject to which they appear (no impression of the self). Since each of these appearings is a fleeting occurrence, inner intuition or experience does not provide anything capable of determining the existence of the subject in time (Allison, p. 299).

That is, on Allison's reading, we have no intuition of the self, hence no knowledge of it, so the self cannot be the *known* permanent required for time-determination.

Now, suppose, as Allison says in this passage, that inner sense provides no intuition of a subject *to which* our representations appear. It could still be the case that inner sense provides us with representations of the self *as it appears*.³ Or, to put the point in Kant's terminology, the fact that we lack intuitions of the transcendental self need not preclude us from having intuitions of the empirical self (see B 155–56, A 492=B 520). Indeed, Kant frequently describes inner sense as the very faculty that provides us with intuitions of the self as an object which appears (see, e.g., B 155, A 342=B 400, A 379, A 33=B 49–A 34=B 50; but cf. B 37).⁴ This is our problem: if Kant holds that the empirical self is knowable through inner sense, why does self-knowledge then fail to provide whatever is required for time-determination?

The problem dissolves if, contrary to the passages just cited, Kant really denies that inner experience gives us knowledge of the (empirical) self. We associate such a negative thesis with Hume rather than with Kant, but Allison would have it that Kant's views about self-knowledge are, in effect, very close to Hume's. If this is correct, Kant might have what he needs to fill the gap in the Refutation.

In the *Treatise*, Hume argued that knowledge (or even the very idea) of a persisting self would require a perpetual impression of the self, but no such impression is to be had:

³ This is not an oversight or confusion on Allison's part; see Note 18.

⁴ Allison describes this position as Kant's "official theory," but he thinks that there is another theory of inner sense at work in Kant's writings (Allison, pp. 258–59).

If any impression gives rise to the idea of self, that impression must continue invariably the same, thro' the whole course of our lives; since self is suppos'd to exist after that manner. But there is no impression constant and invariable. Pain and pleasure, grief and joy, passions and sensations succeed each other, and never all exist at the same time.⁵

Allison's language quoted above ("each of these appearances is a fleeting occurrence") suggests that he may be prepared to ascribe a similar line of thought to Kant. However, if Kant does agree with Hume that the self cannot be directly known as permanent, Kant's reasons for thinking so cannot be the ones Hume adduces—the argument Hume gives would prove too much for Kant's purposes. As Hume notes elsewhere, our experience of outer objects is changing and intermittent, just as inner experience is. For example, you might see an object from one angle, and then from another; you might see it, turn away, and then see it again. By Hume's strictures, such impressions could not constitute knowledge of a persisting outer object,⁶ and knowledge of *any* persisting object (whether inner or outer) is ultimately excluded. According to Kant, however, this would mean that time-determination is completely impossible, contrary to the initial assumption on which the Refutation rests.⁷

To complete the Refutation, Kant needs to establish some disparity between inner and outer sense, such that outer sense gives us direct knowledge of enduring objects, while inner sense does not. Allison's fully developed interpretation is sensitive to this point. On Allison's picture, perceptual experiences afforded by outer sense straightforwardly represent objects in the external world. By contrast, inner sense yields awareness of both perceptual experiences themselves and non-perceptual experiences, such as desires, volitions, and sensations. The former are representations, but of outer objects, while the latter are not representations, or at least not representations of the self. Hence, inner sense provides no representation of the self at all (Allison, pp. 258–59).

Two important sources for Allison's interpretation are a passage from the *Critique of Judgment* and remark by Kant at B 67. The material from the *Critique of Judgment* concerns the status of feelings. When a feeling is regarded as a sensation, Kant says, "it is referred solely to the Subject and is not available for any cognition, not even for that by which the Subject cognizes itself."⁸ Allison reads Kant as here exhaustively dividing the contents of the mind into representations (of outer objects) and items "such as" feelings. Feelings are non-representational and having a feeling is not a cognition of anything (Allison, p. 259).

⁵ Hume, *Treatise of Human Nature* (ed. Selby-Bigge) (Oxford: Clarendon Press, 1967), I.iv.6 (pp. 251–52).

⁶ Hume, *Treatise of Human Nature*, I.iv.2 (pp. 187–218).

⁷ Guyer raises a similar point (Guyer, p. 285).

⁸ Quoted in Allison, p. 259.

I would assign the passage under consideration a somewhat different import. What Kant means by ‘feeling’ (*Gefühl*) is primarily pleasure (*Lust*) or displeasure (*Unlust*).⁹ For Kant, feelings of pleasure and displeasure are something like second-order states (the tendency to remain in or to change one’s first-order state), and they can attach to any representation.¹⁰ The having of such feelings is thus something different from awareness of one’s own inner state as such. The latter, as distinct from feelings of pleasure or displeasure, may be representational and the basis for knowledge—even if feeling is not.

This alternative account gains some support from remarks Kant makes in the *Anthropology*. He there draws a distinction between inner sense proper and interior sense, which is specially concerned with pleasure and displeasure:

The senses, in turn, are divided into the *outer* senses and *inner* sense (*sensus internus*). In outer sense, the human body is affected by physical things; in inner sense, by the mind. We should distinguish between inner sense, which is a mere power of perception (of empirical intuition), and the feeling of pleasure and displeasure—that is, our susceptibility to be determined, by certain ideas, either to hold onto them or to drive them away—which could be called *interior sense* (*sensus interior*).¹¹

Accordingly, Kant’s point in the *Critique of Judgment* may well be that interior sense provides no representations of the self, leaving open the possibility that inner sense does.¹²

Allison’s treatment of inner sense also draws on some comments Kant added to the Transcendental Aesthetic in the B-edition. Kant writes, “the representations of the outer senses constitute the proper material (*eigentlichen Stoff*) with which we occupy our minds” (B 67).¹³ Allison takes this to mean that inner sense has no manifold of its own. For Kant, representation requires a manifold. Given Allison’s reading, then, inner sense (lacking an appropriate manifold) could not provide a representation of the self.

There are several reasons why we ought not to accept this line of interpretation as reflecting Kant’s considered views. First, it is just not clear that Kant’s remark is concerned with the existence of a manifold of inner sense.¹⁴

⁹ See, e.g., the *Critique of Judgment*, p. 16 (Ak. V, 177).

¹⁰ See the *Anthropology*, p. 99 (Ak. VII, 230–31) and the *Critique of Judgment*, p. 139 (Ak. V, 277).

¹¹ *Anthropology*, p. 32 (Ak. VII, 153).

¹² One might not concede, as I have done, that feelings as such carry no representational content (see A 374; also the *Critique of Practical Reason*, Ak. V, 9 n.). Kant’s various conflicting remarks on this general topic are noted by Paton. See H. J. Paton, *Kant’s Metaphysics of Experience* (New York: Humanities Press, 1970), Vol. I, p. 99 n. 4.

¹³ A similar statement is found at B xxxix n.

¹⁴ An alternative reading is to be found in Hans Vaihinger, *Commentar zu Kants Kritik der reinen Vernunft* (New York: Garland Publishing Company, 1976), Vol. II, p. 276.

Moreover, Kant does hold that inner sense has its own *pure* manifold, i.e., time (A 77=B 102), so it is hard to see why or how he would deny that inner sense has an empirical manifold proper to it. Indeed, in several places, he appears to say exactly the opposite.¹⁵

As Allison himself notes, the position he ascribes to Kant is a delicate one (Allison, pp. 259, 265). If inner sense lacks a manifold altogether, it cannot on Kant's terms be the source of any empirical knowledge. That is, we could have no knowledge of our selves as such or of any events in our own mental histories. Such a result would contradict the very assumption of the Refutation that self-knowledge in some form exists. On the other hand, if there is a manifold for inner sense, it can provide knowledge of the self after some fashion. But then why doesn't it provide knowledge of the self as an enduring object?

Allison is well aware of the difficulties here. He maintains that, for Kant, we have self-knowledge in the form of knowledge about our experiences, although we lack knowledge of the self as an object properly speaking:

But since inner experience has no manifold of its own, there are no sensible representations (intuitions) by means of which the self can represent itself to itself as an object. Consequently, in referring its representations to itself in judgments of inner sense, it does not conceive of them as representations of itself in the way outer intuitions are regarded as representations of outer objects. Instead, it conceives of these representations as belonging to itself, as its own "subjective objects." Correlatively, the self regards itself as the substratum or subject in which these representations inhere. Thus, in spite of his theory of judgment, Kant is led to what amounts to a "substratum" or "bare-particular" theory of predication when he deals with judgments of inner sense. (p. 262)

The idea is that, since we do have a manifold for outer objects, we can assign a concrete content to what those objects are. Hence, in predicating a feature of an outer object, we have something more than the concept of a "bare substratum" to serve as the subject of the judgment. By contrast, in a judgment like 'I am having a visual sensation as of a rose', the term 'I' is supposed to lack the requisite concrete content. It follows that such a judgment is not on a

¹⁵ "Inner experience, on the other hand, contains the matter of consciousness and a manifold of empirical inner intuition...." *Anthropology*, p. 22 (Ak. VII, 141–42). See also B 406–7. What the nature of this manifold would be depends upon the scope, function, and mode of operation Kant assigned to inner sense—all difficult and controversial points. A fully adequate treatment of this topic would have to take into account Kant's doctrine of self-affection and his obscure remarks linking inner sense with attention and with the figurative synthesis (see B 153–56). Allison attempts to provide such an extended treatment, but I am doubtful about his claim that self-affection does not produce a manifold for inner sense (Allison, pp. 265–67). For some further considerations against the view that inner sense has no manifold of its own, and for an extended treatment of inner sense in general, see Karl Ameriks, *Kant's Theory of Mind* (Oxford: Clarendon Press, 1982), Chapter VII.

par with judgments of outer sense, and does not constitute knowledge of an object.¹⁶

This account remains deeply problematic. It is hard to see why the self, unlike outer things, has to be represented (if we can talk that way) as a bare substratum rather than as a fully-fledged object. There is a way to characterize or provide content to the notion of oneself: as the subject of one's various experiences, states and mental activities. To be sure, Hume would reject this characterization, but that is because he is hostile to the notion that *anything* can be known as a substance or bearer of properties.¹⁷ Kant, however, doesn't follow Hume in this.

Moreover, it won't do just to stand on the claim that the manifold of outer sense provides no basis for characterizing the self as the subject of various experiences. That would, in effect, leave in place the predicament set out above. The manifold of outer sense, if that is all there is, must function as the basis of some kind of self-knowledge. But if that manifold is plastic enough to provide knowledge of outer experiences as experiences (which Allison allows), there seems to be no reason why it could not provide knowledge of those experiences as states of the empirical self. In turn, there would be no reason to deny that we have direct knowledge of the self as a genuine, persisting object.¹⁸

Clearly, Kant's theory of inner sense involves formidable exegetical and philosophical problems, and, all told, Allison's interpretation may be as good as any other we are likely to get. But just because there are so many problems associated with the doctrine of inner sense, the details of that doctrine are unlikely to provide a clear-cut, fully satisfactory answer to our difficulties with the Refutation of Idealism. At best, it would turn out that the Refutation rests on some obscure and idiosyncratic aspects of Kant's philosophical psychology. Accordingly, we should consider whether Kant might have had some other way to complete the Refutation.

¹⁶ Apparently, though, such a judgment does constitute knowledge of the outer experience itself as what Allison terms a "subjective object." According to Allison, Kant, like Hume, will countenance knowledge of experiences, but not knowledge of the self as such. See also Allison, p. 155.

¹⁷ Hume, *Treatise*, I.iv.5 (p. 232).

¹⁸ Allison offers a further reason to deny that a judgment like 'I am having an experience as of a rose' provides knowledge of the self as an object. He thinks that, in such a judgment, the term 'I' denotes the transcendental self, which on Kant's view is no empirical object (Allison, pp. 262–63). This is an extremely troublesome issue in any case. In support of his construal, Allison refers to R 5453 (Ak. XVIII, 186). In this *Reflexion*, Kant identifies the subject of inner sense judgments as the topic of rational psychology; in a similar passage in the first *Critique*, we find Kant saying: "'I' as thinking am an object of inner sense and called 'soul' ... Accordingly, the expression 'I' as thinking being signifies the object of that psychology which may be entitled 'the rational doctrine of the soul'" (A 342=B 400). I take Kant to be warning us against a confusion of the I known empirically through judgments of inner sense with the transcendental I which isn't known at all—where such confusion is the source of paralogisms.

C. Guyer's Causal Refutation

Paul Guyer has made an extensive study of the whole range of variants on the Refutation Kant considered at one time or other. In Guyer's judgment, the line of thought presented in the B-edition of the first *Critique* was incomplete, and only later did Kant arrive at a more satisfactory anti-skeptical argument.¹⁹ This preferred version (hereafter, "the Causal Refutation") is something of a hybrid between the Refutation proper and a causal argument that Kant originally gave in the *Nova Dilucidatio*. Guyer's proposal has much to recommend it. I will explore it on its own terms; I will also inquire whether Guyer's reconstruction may be taken as completing, instead of replacing, Kant's argument of 1787.

In order to understand and assess the argument Guyer suggests, it will help to examine, briefly, Kant's line of thought in the *Nova Dilucidatio*.²⁰ Kant takes as his point of departure the general principle that if X brings about an alteration in Y, X cannot be identical to Y.²¹ With this principle in hand, he proceeds more or less as follows:

- (N1) If I have existed in the past as well the present, there has been a succession of my states, i.e., I have undergone a transition from my earlier state to my current one.
- (N2) This change in me is the effect of some cause.
- (N3) The cause in question is something distinct from me (by the general principle cited above).
- (N4) Therefore, it follows from the fact that I am determined in time that there must be objects other than myself (and whose existence I can prove by this very argument).

Two observations are in order here. First, this argument is really quite different in conception from the Refutation of Idealism proper. One of the few things clear about the latter is its dependence on the claim that knowl-

¹⁹ See Guyer, p. 285. The principal source for Guyer's argument is Kant's *Reflexionen*, especially R 6313 (Ak. XVIII, p. 615). Guyer is certainly correct that at some point after 1787 Kant considered an anti-skeptical argument that incorporates causal considerations.

²⁰ *Nova Dilucidatio*, pp. 96–98 (Ak. I, 410–11). The importance of this passage was stressed by Moltke Gram in "What Kant Really Did To Idealism," in J. N. Mohanty and Robert Shahan, eds. *Essays on Kant's Critique of Pure Reason* (Norman: University of Oklahoma Press, 1982).

²¹ This causal principle is supposed to follow from the Principle of Sufficient Reason. In the *Nova Dilucidatio*, the causal principle applies only to alterations in simple substances, but Kant also assumes that the soul is simple. Such an assumption would be forbidden in the *Critique* by the results of the Paralogisms. However, Kant does state a version of the causal principle in the *Critique* (A 205=B 250), although now its scope may be restricted to changes in material objects.

edge of a persisting object is necessary for time-determination. By contrast, the *Nova Dilucidatio* argument rests on considerations having to do with causation—not persistence. A second point to note is that the general principle Kant invokes is highly dubious. Imagine that a machine automatically shuts itself off if it detects that it has malfunctioned. In this case, it would seem that the cause of the alteration in the machine is just the machine itself.²² Thus, there are instances in which X causes an alteration in Y, even though X and Y are the same thing, contrary to Kant’s principle. It would be disappointing if Kant could do no better than this.

Guyer’s Causal Refutation isn’t meant to be a simple reworking of the *Nova Dilucidatio* argument. Rather, it is supposed to account for the role of persistence (as well as causation) in time-determination, without making overly strong metaphysical assumptions like the one just pointed out. Whether it succeeds in this will have to be seen.

The argument Guyer presents is complex and multi-layered. For purposes of analysis, I have divided it into seven steps:

(G1) When I am conscious of myself as determined in time, I have to reconstruct my history from the contents of my mental state at the present moment. (p. 309)

(G2) So, consciousness of myself as determined in time involves a complex representational state: I represent both my current state (presumably I am at all times aware of my current state) and my past state (as past). (p. 303) Call the former representation R(NOW) and the latter one R(THEN); call the states they represent NOW (i.e., my current state) and THEN (i.e., my past state), respectively.

(G3) Given that R(NOW) and R(THEN) represent mental states of mine, how do I know that the mental states they represent don’t really co-occur; i.e., how do I know that the mental state represented by R(THEN) really is a *past* state rather than a current one? (p. 307)

(G4) Assume that NOW and THEN are themselves representational states.²³ They can be judged by me now to be successive rather

²² Of course, machines are not simple; see Note 21 immediately above. For an ingenious, but (to my mind) unsatisfactory response to this point see Kant’s *Lectures on Metaphysics* (“Metaphysik L₁”; Ak. XXVIII, p. 267).

²³ Guyer clearly means to restrict the purview of the argument to time-determination of representational mental states, but the motivation for this restriction is unclear. Certainly, the Refutation would turn out to be weaker than Kant apparently intends, if it showed only that *some* consciousness of the self as time-determined (i.e., awareness of the

than co-occurring “only if they are judged to be severally simultaneous with the severally successive states of some enduring object.” (p. 306)

- (G5) If THEN and NOW are judged to be successive by being correlated, respectively, with states S_1 and S_2 of an enduring object, NOW will be numerically distinct from S_2 and THEN will be distinct from S_1 . (p. 308)
- (G6) Moreover, this correlation is achieved by discerning causal relations between S_1 and THEN, as well as between S_2 and NOW. (p. 308)
- (G7) “Just as the successive states of these objects are numerically distinct from the subjective states [they are employed to order], so must the objects whose states they are be numerically distinct from the empirical self on which the objects act.” (pp. 308-9)

The conclusion, of course, would be that knowledge of one’s own history ultimately depends upon the possession of some knowledge of objects other than the self.

One of Kant’s own examples (B 277) provides a useful illustration of how this argument is meant to work. Imagine that, as you now see a sunset, you recall your own past; specifically, you remember having seen the sun overhead hours earlier (G1). In Guyer’s language, you are in a complex representational state, part memory, part not (G2). The question arises, what makes your present recollection of seeing the sun at zenith count as a memory rather than, say, an odd daydream that has nothing to do with the past at all? (G3). You realize that the mental state of seeing the sun overhead was simultaneous with the sun’s being overhead, just as your (present) state of seeing the sun at the horizon is simultaneous with the sun’s being at the horizon. Since the sun can’t be in two different places at once, the sun’s states, and therefore your mental states, must occur at different times. In other words, you can recognize your state of seeing the sun overhead as *past*, in contrast with your present state, by judging your mental states to be correlated with successive states of an enduring object, the sun (G4). In this instance, the sun’s state of being overhead (or of being at the horizon) isn’t identical to your state of seeing the sun overhead (or at the horizon) (G5); instead, the sun’s states cause your mental states (G6). And of course the sun is something distinct from your mind (G7).

history of one’s representational states in particular) requires knowledge of the external world. See also my Note 25, below.

All of this sounds quite reasonable, but the example may make the argument appear stronger than it really is. The capacity to correlate mental states with successive states of an external object might be a sufficient condition for time-determination, but it seems not to be a necessary condition. To see this, suppose that you recollect having felt cold while you are perfectly comfortable. You can, it seems, determine that your state of being cold occurred at some time other than the present, because you cannot be both cold and not cold at the same time. In this case, time-determination would involve the ascription of states only to an enduring self, and not to any other object. To the extent that such a possibility isn't ruled out, Guyer's argument preserves, rather than fills, the apparent gap in the original Refutation.

Guyer would, presumably, respond that the Causal Refutation's conditions for time-determination don't apply here because feeling cold and feeling comfortable aren't representational states. But consider again the example of your remembering having seen the sun overhead, while you currently see the sun on the horizon. Success at time-determination in this instance means recognizing the difference in the temporal locations of two representational mental states. One of these is your (past) state of seeming to see the sun overhead; the other is your more complex (present) state, which involves seeming to see the sun setting while *also* experiencing something like a memory-image of the sun overhead. Since you cannot be in the state of representing *only* the sun overhead while you are also in the state of representing that plus something else, you can judge directly that these states have to be successive rather than co-occurring.²⁴

This objection has been directed against the transition from step (G4) of Guyer's reconstruction to his step (G5). I have tried to show that, in the absence of further argument, there is no reason why time-determination has to proceed by establishing correlations between one's particular mental states and other states numerically distinct from them. Instead, awareness of the mental states themselves might be sufficient. But now what about the remainder of the argument? Why would knowledge that a particular mental

²⁴ Guyer may anticipate this sort of objection. He specifies that time-determination of the self requires a correlation between given mental states and states *distinct* from them only when the mental states can be co-instantiated (p. 307). I am suggesting that in many (if not all) cases of memory Guyer's condition is not met, so quite commonly direct time-determination would be possible. At this point, Guyer might insist upon the need to follow Kant's distinction between logical and real opposition; in Guyer's words, "For at least a large variety of cases, our knowledge that a thing cannot be in two different states at the same time cannot be derived from logical opposition between contradictories, but itself depends upon knowledge of what the causal powers of things are" (p. 253). Perhaps so, but we still seem to be in the position that we might not be able to know *some* things about our mental histories unless we had knowledge of the external world; such an outcome might well leave the problematic idealist unmoved.

state is correlated with some other numerically distinct *state* have to involve knowledge of some *object* distinct from the self?²⁵

There is a Kantian argument, found in the Third Analogy, that to correlate two states in time, i.e., to judge them to be simultaneous, requires positing some kind of causal connection between them. If granted, this yields (G6).²⁶ However, to reach the desired conclusion requires moving from the claim that certain states are causally related to the claim that their respective bearers are numerically distinct. Of course, the general causal principle of the *Nova Dilucidatio* would support this transition to (G7), as Guyer notes (Guyer, p. 308). But, as I indicated earlier, there are liabilities attached to reliance on that principle. It is highly suspect on its own terms; moreover, the principle would make superfluous the argumentation about persistence and time-determination found in both Guyer's reconstruction and in the original Refutation.

Guyer intimates that the principle Kant employs against the problematic idealist isn't quite the metaphysical principle of the *Nova Dilucidatio*, but rather an epistemological analogue of it (Guyer, p. 308). I suppose that such a principle would have to be something like: if X is known (or justifiably believed) to cause a change in Y, then X is distinct from Y. However, this principle seems to have much the same drawbacks as the metaphysical principle. For one thing, it is falsified by the same counterexamples; I might know that a machine caused itself to stop working even though the machine is not distinct from itself. For another thing, the epistemological principle would prove too much. Certainly, I *know* that the cause of my present sensation is the cause of my present sensation. By applying the epistemological principle, I would be able to establish immediately that this cause must be distinct from me, or, in other words, that I am sensing something in the external world.²⁷ This inference would not require any assumptions about the conditions for time-determination, making what Kant explicitly says in the

²⁵ In point of fact, the conclusion follows immediately if total or complete states are involved, because nothing can be in two different total states simultaneously. Even so, this observation might not be sufficiently general for Guyer's purposes.

²⁶ "The co-existence of substances in space cannot be known in experience save on the assumption of their reciprocal interaction" (B 258=A 211); Guyer comments that "objects properly judged to coexist must be known to stand in interaction or a relation of mutual causation" (Guyer, p. 269). However, he himself does not invoke the Third Analogy in this context. One difficulty in doing so would be to understand how causation can be the basis of judgments of temporal succession (according to the Second Analogy), while *mutual* causation serves as the basis for judgments of temporal simultaneity. For an elegant treatment of this issue, see R. I. G. Hughes, "Kant's *Analogies* and the Structure of Objective Time," *Pacific Philosophical Quarterly* 71 (June, 1990). A further problem with invoking the Third Analogy at this point is that the principle of interaction for which Kant argues seems to apply only to objects in space.

²⁷ If my use of something so close to a tautology is disturbing, there are variants like "The cause of my present sensation is causing me to experience the color orange," which also work.

Refutation superfluous. All in all, then, if Kant's own formulation of the Refutation seems too weak, the modifications introduced by Guyer appear to be too strong.²⁸

The criticisms made in this section have been of two kinds. First, I have raised some doubts as to whether Guyer's reconstruction really succeeds in closing the apparent gap in the Refutation of 1787. I have also made the point that Guyer's appeal to causal considerations is difficult to square with the Refutation's emphasis on the need to recognize an object as persisting in time. So, despite Guyer's efforts, it remains open whether Kant had a way to argue, at least in the context of the B-edition Refutation, that knowledge of the self alone is insufficient for time-determination.

D. Conclusion

Allison and Guyer offer two of the most sophisticated treatments of the Refutation of Idealism available in the literature. Each of them takes an initially promising line of interpretation and develops it with care and skill. Finally, however, both commentators fail to explain how Kant would have closed what seems to be a crucial gap in his argument. The limitations of their readings give us reason to pursue alternatives.²⁹

²⁸ In working out the basis for (G4), Guyer invokes the Second Analogy as establishing the need for causal laws as the basis for time-determination, and he combines this with the thesis that mental states are not as such subject to such laws (p. 307). This second thesis raises many difficulties, but let me point out only that these claims would themselves be sufficient to accomplish the aims of the Refutation, again without any considerations pertaining directly to the need for something permanent in time-determination.

²⁹ I hope to explore some such alternatives in future writings. My thanks to Karl Ameriks, Anthony Brueckner, R. I. G. Hughes, and Patricia Kitcher for helpful discussions and correspondence about the Refutation; however, they may very well disagree with the views expressed here. I am also grateful to the anonymous readers for *Philosophy and Phenomenological Research* who provided valuable comments and criticisms.