

Assessing Kant's Master Argument: a review essay on Robert Howell's Kant's Transcendental Deduction⁽¹⁾

Derk Pereboom, University of Vermont

1. Howell's project.

Robert Howell's Kant's Transcendental Deduction is one of the most impressive works ever produced on the Transcendental Deduction in particular, and on Kant's metaphysics and epistemology more generally. Howell's main project is to set out the possible interpretations of the first half of Kant's B-Deduction, distinguishing the possible readings of its premises, and then to evaluate the resulting versions of this argument. The ambiguities derive, for example, from Kant's lack of clarity as to the sort of necessity being invoked and what its scope is to be, the type of self-knowledge at issue in discussions of apperception, and the notion of 'object' at play in claims about knowledge of objects. Ambiguities of these kinds have often been noted, but Howell's work is unique in its rigor in fleshing them out, and in its command of analytical tools in realizing this aim. His overall assessment is negative; in his view every interpretation of the Deduction is beset by insuperable difficulties. The permanent benefit of the Deduction does not lie in the success of Kant's argument, but in the "depth of his insights and the immense fertility of his views," and indeed "the arguments and ideas it suggests ... go as deep, and ramify as widely, as any other arguments and ideas in philosophy" (p. 335).

Howell's book is also one of the most challenging works on Kant ever published. The difficulty of the text does not derive from lack of clarity -- the book is in fact exceptionally lucid. What is daunting to the reader is the complexity of the analysis, which derives largely from Howell's attempt to set out the different sorts of ambiguity in the argument. Following out the ambiguities regarding necessity, for example, demands an understanding of the de re/de dicto contrast, a comprehension of differences in scope, the skill to track the various readings of the premises that these distinctions generate, as well as the patience to assess the multiple versions of the overall argument that result from these distinctions. Despite its intricacy, the book demonstrates that this level of analysis cannot be avoided if one's aim is to evaluate the prospects of Kant's argument. Howell's book is so exceptional at this task that anyone who wishes to tackle the Deduction responsibly must come to terms with it.

All of Howell's versions of the B-Deduction are variations of an overall conception that has been dominant for at least forty years. According to this conception, the first half of the B-Deduction consists of a single argument that begins in §15 and proceeds at least through §20. In rough outline, Kant first argues for the necessary unity of apperception -- that the I think does or can accompany all of my representations (within a single intuition?), and then claims that this unity of apperception is necessary and sufficient for intuitive experience of objects, then that this intuitive experience of objects requires that the unity of apperception be objective, that objective unity of apperception requires judgment, that judgment demands logical forms of judgment, and that logical forms of

judgment applied to intuition are the categories, which legitimately apply to those objects of which we have intuitive experience. In Howell's analysis, insurmountable problems arise at virtually every step. Most crucially, the arguments for the necessary unity of apperception fail, Kant provides no convincing argument that the necessary unity of apperception is sufficient for experience of objects, and the argument that the logical forms of judgment have correlates, the categories, that apply to these objects is inadequate.

In my view, Howell is largely correct in his criticisms of the arguments he finds in the text. But I contend that the structure of the argument is not what he and much of the tradition in the past forty years have claimed it is. I believe that the real structure fares better than Howell's, and that some of his criticisms actually lend support to an alternative picture. The alternative structure cannot save the argument in its entirety. For a case in point, I am convinced by Howell's incisive criticisms of the arguments in §§19 and 20 for the roles that the logical forms of judgment and the categories have in Kant's scheme. Nevertheless, the alternative conception shows that the Deduction has more force and significance than Howell allows. With this theme in mind, I will develop two criticisms of Howell's analysis of the Deduction, focusing first on §16 and then on §17.⁽²⁾ Subsequently, I will raise an issue regarding his treatment of things in themselves.

2. Unity of Apperception.

One of Howell's primary difficulties with the argument of §16 is that Kant does not establish a crucial premise

(S) All of the elements of the manifold of *i* (where *i* is some arbitrary intuition) are such that *H* is or can become conscious, in thought, that all of those elements, taken together, are accompanied by the I think

by contrast with a weaker rival

(W) Each element of the manifold of *i* is such that *H* is or can become conscious, in thought, that the I think accompanies that element. (p. 161)

Howell supposes that if (S) cannot be established, the argument of the Deduction collapses, for only if Kant can demonstrate that the different elements of a manifold together and at the same time are accompanied by the same I think can he apply the combination-cannot-be-given premise of §15, and thus show that *H*'s mind must synthesize these elements. Howell contends that the unity expressed by (W) is insufficient to generate the need for synthesis; rather the co-consciousness expressed by (S) is required. He claims that "if (S) -- or something like (S) -- cannot be established, then the argument of the Deduction, as Kant presents that argument, simply fails" (p. 162). Howell goes on to argue that Kant cannot in fact demonstrate (S), and that hence the argument of the Deduction is in trouble.

I do not believe that Kant needs a premise as strong as (S) -- as opposed to, say, (W), for the argument of §16. In my view, the text bears out the claim that §16 does not crucially involve (S). First of all, I cannot find a passage in this section that unambiguously makes the co-consciousness claim. Furthermore, it is hard to see how (S) would play a role in the argument that one actually finds in §16. A central feature of this argument is Kant's attempt to demonstrate how only a priori synthesis can explain how I might represent the identity of my apperceptive consciousness (B133) or of the apperceiving self or subject (B135) for the different elements of the manifold of intuition to which I can attach the I think. This is a feature of the argument which I believe Howell underemphasizes. In the part of §16 where the crucial argument for a priori synthesis takes place, Kant asks how he might explain how I can "represent to myself the identity of the consciousness in [i.e. throughout] these representations" (B133). The answer comes in the following passage:

That relation comes about, not simply through my accompanying each representation with consciousness, but only in so far as I conjoin one representation with another, and am conscious of the synthesis of them. Only so far, therefore, as I can unite a manifold of given representations in one consciousness, is it possible for me to represent to myself the identity of the consciousness in [i.e. throughout] these representations. (B133)

Missing in Howell's account of the argument is an explanation as to why uniting the manifold of representations in one consciousness explains my representing, specifically, the identity of the apperceptive consciousness of the different representations to which I can attach the I think, or equivalently, the identity of the apperceiving I or subject for the different representations to which I can attach the I think.

One might envision several types of explanation for my representing this identity of the apperceiving subject. Kant seems not to believe that the representation of the identity in question is basic, and thus has no ground or explanation, or that it is founded solely on my ability to attach the (same or at least same-sounding) I think to the different states. Another possible explanation is introspective, and depends on conceiving introspection as analogous to external perception. On this proposal, I have inner intuitions that enable me to represent this identity for each of my various self-attributions, and the way I represent the identity of the subject is similar to the way I represent the identity over time of ordinary objects -- for instance, by noting similarities in the contents of different intuitions of them. But Kant and Hume would concur that this is not the way I represent the identity of the apperceiving subject, since both would agree that I do not have an intuition of it. The second type of explanation, which in my view Kant endorses, is that I have a more indirect way of representing this identity. As Henry Allison argues, Kant claims that I cannot represent this identity by directly intuiting it, so instead, this representation must instead depend on my apprehending a feature of my representations.⁽³⁾ Kant contends that the relevant feature is some type of unity of these states. If the representations I can attribute to myself possess a unity of the appropriate type, and if I apprehend this unity, then I can represent the apperceiving subject of any one of them as identical with that of any other.

What sort of unity must I recognize among my representations in order to explain how I might represent this identity? Here Howell's reading might be brought in: the requisite unity among representations is that they simultaneously be objects of conscious awareness.⁽⁴⁾ But what might motivate this move? Note that I represent the subject as identical for self-attributed representations that are not co-conscious, so co-consciousness could not explain generally how I represent this sort of identity. A plausible account is that the unity consists in certain intimate ways in which representations in a single subject are usually related. One candidate is that a single subject's representations are inferentially integrated to a high degree, much more so than are representations across discrete subjects. Another is that the representations of a single subject can and often do causally interact with each other without a non-mental intermediary. On this interpretation, then, my recognition that my representations have these intimate relations accounts for my capacity to represent the identity of the apperceiving subject for different self-attributions. On Kant's view, a priori synthesis would be required for my production and recognition of these intimate relations among my mental states. I do not see how awareness of the unity expressed by (S) would provide a better explanation of my ability to represent the identity of the apperceiving subject than would the unity expressed by (W). Nor, for that matter, do I see a reason for thinking that the unity expressed by (S) provides a better foundation for the universal applicability of the categories than does the unity expressed by (W). Because (W) is relevantly weaker than (S), Kant's argument can at this point survive Howell's criticism.

3. The structure of the B-Deduction.

An obstacle to Howell's claims about the structure of the first half of the B-Deduction is that Kant's summary of this argument in §20 begins with a premise from §17, and not with a premise from §15 or §16. If the crucial premise of the argument is to be found in §15 or §16, as it is in Howell's view, one would expect Kant to highlight this fact in his summary. Notice, however, that the summary of §20 has a correlate in the argument from below in A, which aims to demonstrate that synthesis by means of the categories is necessary for explaining our representations of objects. The argument from below contrasts with the argument from above, which is designed to show that this type of synthesis is required to explain pure apperceptive self-consciousness (A119). I contend that §§17-20 is essentially an argument from below, and that it functions as Kant's main argument in the B-Deduction. §16 contains a truncated version of the argument from above, whose function in the B-Deduction is to explain the nature of the necessary unity of apperception, which according to §§17-20 is necessary but not sufficient for experience of objects. It is truncated since although it does claim the need for synthesis, it does not proceed all the way to the applicability of the categories.

Howell maintains that §17 crucially contains a demonstration of the representation of objects from premises about the theory of self-consciousness developed in §§15-16. The crucial claim for this interpretation is that the transcendental unity of apperception is not only a necessary but also a sufficient condition for the representation of objects. In my interpretation, by contrast, the transcendental unity, and more precisely, the synthesis that

explains our consciousness of the identity of the subject, is only a necessary condition for the representation of objects, and, moreover, a condition which Kant only proposes in §17, and does not aim to demonstrate until §§18-20. One might be attracted to Howell's interpretation because it would have Kant demonstrating the existence of objects (in some sense) on the grounds of certain premises about self-consciousness. The Deduction would then not only aim to undermine Humean scepticism about a priori concepts, but also a kind Cartesian skepticism about the existence of objects. Attractive as this may seem, this is not an explicit aim of the Deduction, and neither has anyone ever found even a promising argument against Cartesian skepticism in the B-Deduction itself.

In Howell's conception, the argument from the transcendental unity of apperception for the representation of objects is compressed into this single paragraph (pp. 226-7):

(A) Understanding is... the faculty of knowledge. This knowledge consists in the determinate relation of given representations to an object. An object, however, is that in the concept of which the manifold of a given intuition is united. Now all unification of representations demands unity of consciousness in the synthesis of them. Consequently it is the unity of consciousness that alone constitutes the relation of representations to an object, and therefore their objective validity and the fact that they are modes of knowledge and upon it therefore rests the very possibility of the understanding. (B137)

Allison, who agrees with Howell that this passage aims to show that the unity of consciousness is not only necessary but also sufficient for experience of objects, raises an objection for this interpretation. In his view and in Howell's (p. 227), the claim for sufficiency is found in this sentence in particular:

(1) it is the unity of consciousness that alone constitutes [ausmacht] the relation of representations to an object, and therefore their objective validity...⁽⁵⁾

and Kant presents (1) as a direct consequence of the premise that

(2) all unification of representations demands unity of consciousness in the synthesis of them.

But given this picture, Allison points out that Kant is apparently guilty of a non sequitur, since (2) would clearly license only the conclusion that the unity of consciousness is a necessary condition for the representation of an object, and not also a sufficient condition. Howell voices a similar worry: "In 17 Kant simply does not make this inference clear, and an air of blatant fallacy hovers over this part of his reasoning." (p. 228)

Allison's way out of this problem is to suggest that in §17 Kant employs a minimal sense of 'object,' according to which an object is the result of any synthetic process of unification.⁽⁶⁾ Howell disavows Allison's move, and to my mind presents decisive reasons to do so (p. 366, n. 39). Instead, Howell claims that Kant's best strategy is to establish that the holding of "the unity of apperception can be shown to be the source of, and so to

imply," the representation of objects, although in Howell's view that argument also ultimately fails (pp. 228ff). I agree that for Kant (the synthesis that produces) the unity of consciousness (in turn) produces or is the source of one's representation of objects. But this is not clearly equivalent to claiming that the unity of consciousness by itself is sufficient for representation of objects. By analogy, we might legitimately say that the process of smelting and molding steel is the source of steel girders, while strictly speaking these processes are not all by themselves sufficient for the production of steel girders -- for instance, particular sorts of molds may also be required. Admittedly, we sometimes use 'sufficient for' in a loose way, allowing a condition C to be sufficient for D when strictly speaking it is not, but only contextually understood background conditions need to be supplied in addition. But in a context of an anti-skeptical argument in which one is attempting to derive, say, the existence of objects from a minimal premises, one cannot use 'sufficient' quite so loosely.

More generally, sufficiency is a tight logical relation, and the fairly loose terminology Kant uses in this discussion does not force the sufficiency reading on us. Allison and Howell both venture that Kant's sentence

(1) Consequently, it is the unity of consciousness that alone constitutes [ausmacht] the relation of representations to an object, and therefore their objective validity...

should be read as a statement of the sufficiency claim. Look again at the paragraph (A) of which (1) is a component. What Kant is claiming there is that cognitions of objects consist in some determinate relation of representations to objects, and as (1) indicates, this relation is constituted or produced by a synthesis that crucially involves the unity of consciousness. But (1) does not entail that the synthesis that involves unity of consciousness cannot take place without being the source of a relation of representations to objects. Turning to our analogy, the smelting and molding of steel are processes that constitute or produce steel girders, but it does not follow that the processes of smelting and molding steel cannot take place without the production of steel girders. Just as the production of steel girders requires certain molds in addition to the processes of smelting and molding steel, so the production of representations of objects may well require particular concepts of objects in addition to the synthesis that involves the unity of consciousness.

None of this is to reject what is arguably Kant's picture, that the synthesis that produces the unity of apperception also produces all representations of objects. This idea might well be the sense of his claims that "the synthetic unity of apperception is the highest point to which we must ascribe all employment of the understanding" (B134 note), and "the principle of the synthetic unity is the supreme principle of all employment of the understanding" (B136, title).⁽⁷⁾ But this idea does not force the sufficiency condition on us. These texts are consistent with the notion that the synthesis that accounts for the unity of apperception is the most significant element in the production of representation of objects, and this is how I read them. At the same time I do acknowledge that in the A-Deduction (for example, at A108) Kant is more plausibly interpreted as endorsing the sufficiency condition.⁽⁸⁾ But I nevertheless contend that establishing such a thesis is not

an element in the argument of the B-Deduction, and in particular, not a part of the project of §17.

Suppose that Howell is right that Kant's argument for sufficiency in §17 is only scarcely discernible and in any event unsuccessful. Add to this that nowhere in §17 does Kant unambiguously state the sufficiency claim, and that in the summary of §20, it plays no role. The conclusion that this thesis has no part in Kant's argument becomes quite plausible. How might §17 function in the argument absent the sufficiency claim? In my view, the main role of this section is to provide a characterization of a representation of an object that incorporates a challenge to Humean associationism and thereby initiates the argument from below. Kant proposes that an object is "that in the concept of which a manifold of a given intuition is united" (B137). The type of unification of a manifold he refers to in this characterization is synthetic, and in particular it requires the synthesis that produces the unity of apperception, for immediately following the characterization of an object he claims that "all unification of representations demands unity of consciousness in the synthesis of them" (B137). This depiction is designed to present Kant's view about the processing required to produce a representation of an object, and hence, one intended to provide a challenge to Humean associationism. Most significantly, I think this characterization does not set out a position Kant expects us to accept without argument, but rather, one he aims to establish in §§18-20. In §18 in particular, Kant's argument proceeds by adducing uncontroversial features of our representations of objects, and then showing that his theory of a priori synthesis is needed to account for those features. To my mind, this is a promising argument, and for this reason the Deduction as a whole fares better given the structure I have just outlined than it does on a reading that involves the implausible sufficiency claim.

4. Things in themselves.

Howell believes that a serious difficulty confronts Kant's position on things in themselves and appearances, no matter how that view is taken. He first distinguishes between two interpretations of Kant's view. Appearing theories

take the object of knowledge to have (or at least to be able to be conceived to have) an existence in itself. They take that object then itself to be grasped by and known by the mind as that object itself appears to the mind via the operations of the mind's cognitive apparatus... whatever form the object takes as it appears and however its appearing comes about, it is that object itself that appears to -- and it is that object itself that is grasped and known by -- the mind. It is not simply some idea, representation, or other mental simulacrum of the object which the mind grasps and knows (pp. 37-38.)

Appearance theories also take various objects to exist in themselves. However, unlike appearing theories,

appearance theories deny that those objects themselves appear to the mind (or are grasped by or known by the mind) in any form. Rather, those objects generate appearances of

themselves -- mental entities like ideas, sensations, or sense data -- in the mind. These appearances are totally distinct from the objects. And it is these appearances which the mind knows, and not the objects themselves. (p. 38)

Howell points out that for each of these theories one can find passages that bear its imprint. In his view, however, there is a serious problem that undermines both of them. Consider first the appearing theory. Suppose that as a result of having an intuition someone knows that the tree in front of him is conical, and so

(P) H knows that the tree is conical

But our states of knowledge have an existence in themselves, and hence (P) is true in the world W of objects as they exist in themselves. At the same time, it is a familiar and widely accepted that

(Q) If H knows that p, then p.

From this it follows that at W

(R) The tree is conical.

is true. The tree referred to in (P) is, by the appearing theory, strictly identical to the tree referred to in (R). But "if an entity that occurs in a world U can be truly said, at U, to have a certain property, then that entity has that property in its occurrence in U" (p. 43), and thus (R) entails that the tree, as it occurs in W, is conical. This claim contravenes Kant's doctrine that as they exist in themselves objects do not have spatial properties. Furthermore, from (P) and the above reasoning it follows that there is an object that in its occurrence in W is such that H knows it to be a conical tree. This result is inconsistent with Kant's claim that objects, as they exist in themselves, are unknown by us.

Howell claims that the appearance theory falls by a similar argument, since on this view it is also true that "intuitions have an existence in themselves in their roles in our knowledge." From this it follows that "the appearances that we thereby know also have such an existence." Howell concludes that if Kant "adheres to the appearance theory version of his account of knowledge, he arrives at an inconsistency similar to the inconsistency that we have just derived from his appearing theory." (p. 45).

Despite the existence of texts that reflect an appearing theory, I suspect that a version of the appearance theory fits best with Kant's overall position. Furthermore, I believe that Howell is mistaken in his claim that the appearance theory falls to the same sort of argument that undermines the appearing theory. I agree with Howell that intuitions, and our mental states in general, must have an existence in themselves. But this fact does not commit Kant to the claim that on the appearance theory, the appearances that we know by our intuitions will have an existence in themselves. In the best version of the appearance theory with which I'm acquainted appearances are intentional objects of intuitions. A view of this type has been articulated by Robert Adams,⁽⁹⁾ and in fact, Howell himself

spells out a version of this position in his discussion of the appearing theory (pp. 47-9). In one important respect an analogy can be drawn between appearances and hallucinatory objects. By our ordinary notions about the real world, even though the hallucination itself exists in the real world, the pink elephant that is hallucinated does not exist the real world but only in the Hallucinatory World. In the real world the pink elephant does exist as an intentional object of a hallucination, but intuitively this fact does not entail that the pink elephant itself exists in the real world. Similarly, even though the intuition of the conical tree exists in W, the conical tree itself does not exist in W but only in the phenomenal world. In W the conical tree does exist as an intentional object of an intuition, but this does not entail that the conical tree itself exists in W. Hence, what R's being true at W amounts to is that at W the tree -- understood as an object that exists only in the phenomenal world and not in W -- is conical. This appearance theory avoids any conflict with Kant's claims that as they exist in themselves objects do not have spatial properties, and that as they exist in themselves, objects are unknown by us.

Now when Howell discusses an option of this sort in the context of appearing theories, he points out that it is inadequate as an appearing theory because it "denies that there is an identity between an object that in its occurrence in the phenomenal world W' is a conical tree and the object that in its occurrence in W is a nonspatiotemporal, unknowable thing" (p. 48). But this is not to say that this option is inadequate as an appearance theory. To be sure, appearance theories in general do not give Kant the identity Howell underscores, and the texts provide evidence that Kant wants that identity. Moreover, on Howell's reading, the identity is very important to Kant's project. But I suspect, first of all, that this appearance theory gives Kant as much as he can have given his variety of transcendental idealism. Howell might well concur here (pp. 48-50). But I also think that notion that there is such an identity -- an identity between objects that in their occurrence in the phenomenal world have spatiotemporal properties and objects that in their occurrence in the world W of things in themselves are nonspatiotemporal, unknowable things -- is of little importance to Kant's philosophical project, and that if he had simply avoided making the remarks that suggest it, he would not have been in trouble. Howell would surely disagree here, but that discussion will have to take place elsewhere.

Kant's Transcendental Deduction is a defining work in this area of study, and anyone currently working on the topic must understand the interpretation and critique that Howell develops. Although it is one of the most challenging books on Kant, seeking out its insights is indeed richly rewarding.⁽¹⁰⁾

Notes

1. Robert Howell, Kant's Transcendental Deduction (Dordrecht: Kluwer, 1992). References to Kant's Critique of Pure Reason are from the translation by Norman Kemp-Smith, (New York: St. Martin's Press, 1965).
2. See my "Self-Understanding in Kant's Transcendental Deduction," Synthese 103 (1995), pp. 1-42.

3. Henry Allison, Kant's Transcendental Idealism (New Haven: Yale University Press, 1983), pp. 142-4; see also Paul Guyer, Kant and the Claims of Knowledge (Cambridge: Cambridge University Press, 1987), pp. 133-139.

4. Allison, p. 143.

5. Allison, p. 145. Richard Aquila also argues for the sufficiency claim; see his Matter and Mind (Bloomington: Indiana University Press, 1989), pp. 159ff..

6. Allison, p. 146; cf. pp. 27, 118-9, 135.

7. Howell emphasizes these passages in personal correspondence.

8. 8. At A108 Kant says:

The original and necessary consciousness of the identity of the self is thus at the same time a consciousness of an equally necessary unity of the synthesis of all appearances according to concepts, that is, according to rules, which not only make them necessarily reproducible but also in so doing determine an object (Gegenstand) for their intuition, that is, the concept of something wherein they are necessarily interconnected..

9. Robert Adams, Leibniz (New York: Oxford University Press, 1994), pp. 216-28.

10. Thanks to Robert Howell for excellent comments on a draft of this essay.