



# Stanford Encyclopedia of Philosophy

## Kant's Transcendental Arguments

*First published Fri Aug 21, 2009; substantive revision Wed Dec 11, 2013*

Among Immanuel Kant's (1724–1804) most influential contributions to philosophy is his development of the transcendental argument. In Kant's conception, an argument of this kind begins with a compelling premise about our thought, experience, or knowledge, and then reasons to a conclusion that is a substantive and unobvious presupposition and necessary condition of this premise. The crucial steps in this reasoning are claims to the effect that a subconclusion or conclusion is a presupposition and necessary condition of a premise. Such a necessary condition might be a logically necessary condition, but often in Kant's transcendental arguments the condition is necessary in the sense that it is the only possible explanation for the premise, whereupon the necessity might be weaker than logical. Typically, this reasoning is intended to be *a priori* in some sense, either strict (Smit 1999) or more relaxed (Philip Kitcher 1981, Pereboom 1990). The conclusion of the argument is often directed against skepticism of some sort. For example, Kant's Transcendental Deduction targets Humean skepticism about the applicability of *a priori* metaphysical concepts, and his Refutation of Idealism takes aim at skepticism about external objects. These two transcendental arguments are found in the *Critique of Pure Reason* (1781, 1787), but such arguments are found throughout Kant's works, for example in the *Critique of Practical Reason* (1788), in the *Critique of the Power of Judgment* (1790), and in the *Opus Posthumum* (Förster 1989). This article focuses on the Transcendental Deduction, the Refutation of Idealism, and more recent transcendental arguments such as P. F. Strawson's (1966) that are inspired by these two examples.

- [1. The Transcendental Deduction](#)
    - [1.1 Apperception and its Unity](#)
    - [1.2 Representations of Objects and the Single-Substantive-Premise Interpretation of the B-Deduction](#)
    - [1.3 Universality and Necessity in our Experience of Objects](#)
    - [1.4 Logical Forms of Judgment and the Categories](#)
    - [1.5 The Second Step of the B-Deduction](#)
  - [2. The Refutation of Idealism](#)
  - [3. Contemporary Kantian Transcendental Arguments](#)
  - [4. Concluding Remarks](#)
  - [Bibliography](#)
    - [Primary Literature](#)
    - [Secondary Literature](#)
  - [Academic Tools](#)
  - [Other Internet Resources](#)
  - [Related Entries](#)
-

# 1. The Transcendental Deduction

The Transcendental Deduction (A84–130, B116–169) is Kant's attempt to demonstrate against empiricist psychological theory that certain *a priori* concepts correctly apply to objects featured in our experience. Dieter Henrich (1989) points out that Kant's use of '*Deduktion*' redeploys German legal vocabulary; in Holy Roman Empire Law, '*Deduktion*' signifies an argument intended to yield a historical justification for the legitimacy of a property claim. In Kant's derivative epistemological sense, a deduction is an argument that aims to justify the use of a concept, one that demonstrates that the concept correctly applies to objects. For Kant a concept is *a priori* just in case its source is the understanding of the subject and not sensory experience (A80/B106; Strawson 1966: 86). The specific *a priori* concepts whose applicability to objects of experience Kant aims to vindicate in the Transcendental Deduction are given in his Table of Categories (A80/B106); they are Unity, Plurality, and Totality (the Categories of Quantity); Reality, Negation, and Limitation (the Categories of Quality); Inherence and Subsistence, Causality and Dependence, and Community (the Categories of Relation), and Possibility-Impossibility, Existence-Nonexistence, Necessity-Contingency (the Categories of Modality).

David Hume in effect denies that a deduction can be provided for a number of metaphysical concepts — *ideas*, in his terminology — including the ideas of personal identity, of identity over time more generally, of the self as a subject distinct from its perceptions, and of causal power or force (1739, 1748). In Hume's view, a concept can only be validated by finding a sensory experience, that is, an *impression*, in particular the one that is the 'original' of that idea, which must resemble that idea. But because, for example, any attempt to find an impression of causal power turns out to be fruitless, Hume concludes that this idea does not legitimately apply (1748: §7). In Kant's terminology, Hume is determining whether one might provide an *empirical* deduction of the concept of causal power (A85/B117), and from the failure of the attempt to do so, he concludes that this concept lacks *objective validity*, that is, it does not apply to the objects of our experience.

Hume's position on the deducibility of *a priori* metaphysical concepts is Kant's quarry in the Transcendental Deduction. But Kant concurs with Hume's proposal that no empirical deduction can be supplied for such concepts. Instead, he sets out to provide a different sort of justification for their use, one that is *transcendental* rather than empirical. Such a transcendental deduction begins with a premise about any possible human experience, a premise to which reasonable participants in the debate can be expected initially to agree, and then contends that a presupposition and necessary condition of the truth of that premise is the applicability of the *a priori* concepts in question to the objects of experience. Kant's Transcendental Deduction features a number of component transcendental arguments. Each begins with a premise either about the self-attributability of mental items, *apperception*, or else a premise about the necessity and universality of some feature of our experience of objects. Kant strategy is then to establish a theory of mental processing, *synthesis*, by arguing that its truth is a necessary condition for the truth of such a premise, and then to show that the *a priori* concepts at issue -- the categories -- have an essential role in this sort of mental processing. On a metaphysical idealist interpretation of his position, the objects of experience result from this mental processing, and it is due to the role that the categories have in this processing that they correctly apply to objects. In the Transcendental Deduction Kant would thus intend to secure a normative claim, that the categories correctly apply to the objects of our experience, by establishing a

psychological theory (Henrich 1989; Patricia Kitcher 1990: 2–29; for information on related argumentative strategies in Kant's historical context, see Kuehn 1997 and Dyck 2011). The Transcendental Deduction presents general considerations supporting the applicability of all the categories to the objects of our experience; it does not concentrate on the applicability of specific categories. This more focused task is taken up in various sections of the *Analytic of Principles* (A130–235/B169–287).

In the *Metaphysical Deduction* (A66–83, B92–116) Kant intends to derive the categories from the specific modes or forms of any human thought about the world, the *logical forms of judgment*. The *Metaphysical Deduction* has an essential role to play in the *Transcendental Deduction*, and we will discuss this argument at an appropriate juncture (when we reach §19 of the B-Deduction).

For Kant, the most significant rival theory of mental processing is that of his target, Hume. Hume concurs that a theory of experience requires an account of the processing of mental items, but he denies that such an account demands *a priori* concepts or issues in their legitimate applicability to experience. According to his theory, *associationism*, our mental repertoire consists solely of perceptions, all of which are sensory items — the more vivid impressions, which constitute sensory experience, and their less vivid copies, the ideas, which function in imagination, memory, reasoning, and conceptualization (1748, §§2, 3). Association is the process by which these perceptions are related and ordered. A signature feature of association on Hume's view is that it requires no resources distinct from the perceptions themselves. How perceptions are ordered is solely a function of what perceptions alone can provide. Significantly, a subject not constituted solely of perceptions has no role in Hume's theory; the Humean subject is just a collection of perceptions (1739: I, IV, vi). These last two features in particular make Humean associationism a highly economical theory, which provides it with an initial advantage over Kant's more complex view. However, Kant contends that associationism cannot account for the facts to which the premises of the *Transcendental Deduction* appeal, and that synthesis by *a priori* concepts, that is, the categories, is required in addition.

Kant characterizes synthesis as “the act of putting different representations together, and grasping what is manifold in them in one cognition” (A77/B103); it is a process that “gathers the elements for cognition, and unites them to form a certain content” (A78/B103). Synthesis takes multiple representations — in Kant's terminology, a ‘manifold’ — and connects them with one another to produce a single further representation with cognitive content (Patricia Kitcher 1990, 2011). This process employs concepts as modes or ways of ordering representations. A claim critical to the *Transcendental Deduction* is that it is the categories by means of which manifolds of our representations are synthesized. Since the understanding of the subject is the source of the categories, and also a faculty that yields synthesis, the subject plays a crucial role in mental processing. It is important for Kant's view on mental processing that this subject is distinct from its representations, and this is another respect in which it differs from Hume's theory.

This discussion will focus on the *Transcendental Deduction* in the second edition of the *Critique of Pure Reason* (1787) — the *B-Deduction* — thoroughly rewritten and rethought relative to the *A-Deduction* of the first edition (1781). On my reading, in §§16–20 of the B-Deduction Kant employs a two-pronged strategy for defeating associationism and establishing synthesis. The first, contained in §§15–16, is designed to show that association cannot account for an aspect of one's consciousness of one's self that Kant refers to as *the consciousness of its unity*, and that synthesis is required to provide

this explanation. This type of argument he calls an *argument from above*, signifying that it begins with a premise about self-consciousness. Correlatively, in §§17–20 we find an *argument from below*, by which Kant intends to establish that synthesis by means of the categories is needed as a necessary condition of how we represent objects (the above/below terminology is derived from A119; for relevant historical background, see Carl 1989, 1992). On several other readings, as we shall see, the B-Deduction is a more unified argument. There are reasons to accept and reasons to resist interpretations of this kind.

## 1.1 Apperception and its Unity

The argument from above in §16 can be divided into two stages. The goal of the first is to establish the various components of *the principle of the necessary unity of apperception*. The second stage aims to show that synthesis is a necessary condition for the aforementioned aspect of self-consciousness, which this principle highlights. *Apperception* is the apprehension of a mental state – a representation, in Kant's terminology – as one's own; one might characterize it as the self-ascription or self-attribution of a mental state (Strawson 1966: 93–4). In Kant's conception, my apperception has *necessary unity* since all of my representations *must* be grounded “in pure apperception, that is, in the thoroughgoing *identity* of the self in all possible representations” (B131–2, emphasis mine). By this he means that:

(*The principle of the necessary unity of apperception*) It must be the case that each of my representations is such that I can attribute it to my self, a subject which is the same for all of my self-attributions, which is distinct from its representations, and which can be conscious of its representations (A116, B131–2, B134–5).

One might note three aspects of the meaning of this principle. First, as pointed out earlier, Kant maintains that the apperceiving subject is not itself a collection of representations. Kant affirms that I have no inner intuition of the subject (e.g. B157), and this claim would conflict with the subject's being a collection of representations, since he holds that I can intuit my representations by inner sense (e.g. A33/B49). Also, in §16 Kant remarks: “through the ‘I’, as simple representation, nothing manifold is given; only in intuition, which is distinct from the ‘I’, can a manifold be given” (B135). If he believed the apperceiving subject to be a collection of representations, it would be surprising for him to deny that anything manifold is given through my representation of it. Second, my ability to attribute representations to myself as subject of them is *pure*, as opposed to *empirical* apperception. This means that I have this ability not in virtue of Humean inner perception, or Kantian inner intuition, but rather independently of any such empirical faculties. However, Kant repeatedly affirms that the purity of this apperception does not imply that the subject to which one's representations can be attributed is intuited — represented as an object — in a purely rational or *a priori* way (e.g., B406–9). Third, Kant states that pure apperception is *original*, and the explication he provides is that “it is that self-consciousness which, while generating the representation ‘I think’ ... cannot itself be accompanied by any further representation” (B132). I cannot have an intuition or any other type of representation of myself as apperceiving subject other than by way of ‘I think...’-type thoughts, and hence these thoughts are the original representations of this subject (e.g., A350). At the same time, by virtue of my capacity for apperception, I can have a kind of propositional grasp of the apperceiving subject; Kant affirms that in apperception, I am conscious *that* I exist as subject (B157).

Kant initiates the first stage of the argument in §16 by claiming:

It must be possible for the 'I think' to accompany all my representations; for otherwise something would be represented in me which could not be thought at all, and that is equivalent to saying that the representation would be impossible, or at least would be nothing to me. (B131–32)

On one interpretation, the sense in which a representation would be impossible or nothing to me if it could not be accompanied with the 'I think' is simply that I could not then become conscious of it (Guyer 1987: 139–44). It is credible that for any representation of which I am conscious, I can attribute it to myself as subject, assuming my mental faculties are in working order, and if no controversial account of the nature of the subject is presupposed. However, the claim that I can become conscious of *each* of my representations, and that it is therefore possible for me to attribute each of them to myself as their subject, is likely false. Plausibly, some of my representations are so thoroughly subconscious that I cannot attribute them to myself, while they are nevertheless mine due to the causal relations they bear to other representations and to actions that are paradigmatically mine. Fortunately, however, the premise that each of my representations is such that I can attribute it to myself is not crucial for the argument from above. Rather, the premise Kant ultimately singles out is less committed, and focuses more specifically on the identity or sameness of the subject of different self-attributions, and my being conscious of this identity.

A number of interpreters, including Robert Howell (1992) and James van Cleve (1999), maintain that the argument of §16 requires a premise affirming the possibility of my being *simultaneously* conscious of the multiple elements of my representations, and that the implausibility of this premise jeopardizes the soundness of the argument. Howell's specific objection is that Kant does not establish the following 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*.

(S) asserts that all of the individual elements of the selected intuition are such that the subject is or can become conscious of them simultaneously. In Howell's interpretation, only if the elements of an intuition together and at the same time are accompanied by the same *I think* will it be plausible that the subject must synthesize these elements (Howell 1992: 162). On van Cleve's reading, it is required that for any intuition that I have, I actually become simultaneously conscious of its elements. If this co-consciousness were just merely possible, Kant could only conclude that the resulting representation is only possibly subject to the categories (van Cleve 1999: 84). On an interpretation of this type, the mechanism of the argument of §16 is to adduce a kind of unity or combination that my representations actually exhibit, and then to argue that this unity requires synthesis by means of the categories as a condition or for its explanation. Kant is thus read as contending that actual co-consciousness is a type of unity that demands synthesis by means of the categories, and that any variety of unity short of co-consciousness will be inadequate to establishing this objective of the Deduction.

Howell points out that (S) contrasts with a weaker claim:

(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. (Howell 1992: 161)

which allows that the individual elements of the intuition are such that the subject can only become conscious of each separately, perhaps in turn. He contends, however, that the unity expressed by (W) is insufficient to generate this need for synthesis. If I am merely (possibly or actually) serially conscious of the elements of an intuition, it won't be required that I synthesize them into a unified intuition. Howell goes on to argue that while (W) is credible, Kant cannot in fact establish (S); it is implausible that such co-consciousness for any arbitrary intuition is actual or even genuinely possible for us. Consequently, the soundness of this argument, and the overall argument of the B-Deduction is imperiled.

However, it may be that only the weaker premise (W) is required for the success of the argument of §16. First, although the co-consciousness premise (S) might be *suggested* in §16 by the following text:

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 (B133)

that claim is not uncontroversially made here. For by the assertion that I “am conscious of the synthesis of them,” i.e., of these representations, Kant may mean only that I am conscious that they stand in a certain intimate relation to one another, for instance, that that they are integrated with each other in a way distinct from how mine are integrated with yours, which does not require co-consciousness (Pereboom 1995).

Moreover, it might well be that the argument of §16 features a subtlety that obviates the need for actual co-consciousness. The text indicates that the argument crucially turns on the claim that only *a priori* synthesis – that is, synthesis by *a priori* concepts – can explain *how I might represent the identity of my apperceptive consciousness* (B133) or *how I might represent the identity of the apperceiving subject* (B135) for different elements of the manifold of intuition to which I can attach the *I think*. The inadequacy Kant claims for “empirical consciousness,” that is, for consciousness according to Humean psychological theory, is that “it is in itself dispersed (*an sich zerstreut*) and without relation to the identity of the subject (*und ohne Beziehung auf die Identität des Subjects*)” (B133). One implication of this passage is that Hume's theory does not have the resources needed to explain how I can self-attribute various of my representations or their elements to a subject that is both conscious of them and the same subject for each act of their self-attribution. This objection does not beg the question in Kant's dispute with Hume, since it assumes only a claim that uncommitted and reasonable participants in this discussion would not want to initially deny, that the conscious subject of these apperceptive self-attributions is the same. Hume's theory lacks the resources to account for this identity. Hume himself provides no account of apperception, but possibilities for a Humean account are that apperceptive consciousness amounts to perceptions that are intrinsically self-conscious, or else consists in perceptions of perceptions. But intrinsically self-conscious perceptions would be distinct from one another, as would perceptions of perceptions; and thus they too would be “dispersed” (B133), and share no common element. Hume might propose to explain our sense of the identity of the conscious subject of different self-attributions by the intrinsically self-conscious perceptions or the perceptions of perceptions being components of a single causally coherent bundle. However, this bundle would not itself be conscious of perceptions. Consciousness of

perceptions would instead be an intrinsic feature of an individual self-conscious perception or a feature of individual perceptions of a perception. In Kant's conception, by contrast, accounting for our sense of the identity of the conscious subject of different self-attributions requires that this subject be distinct from its representations.

The second stage of the argument of §16 highlights another implication of the claim that “the empirical consciousness, which accompanies different representations, is dispersed and without relation to the identity of the subject” (B133): that Hume's theory lacks the resources to account for my *representation-relation* to the identity of the subject; that is, this view cannot explain how I can “represent to myself the *identity of the consciousness in [i.e. throughout] these representations*” (B133) (for an opposing view see Dickerson 2004: 95-8). We might imagine several kinds of explanation for my representation of this identity. One candidate is that inner sense allows me to represent this identity: the way I represent the sameness of the subject is akin to how I commonly represent the identity over time of ordinary objects — by a cognitive sensitivity to similarities among the intrinsic properties represented. However, Kant and Hume concur that this is not how I might represent the identity of the apperceiving subject, since they agree that by inner sense I cannot represent any intrinsic properties of such a subject. A second kind of explanation, which Kant endorses, is that I have an indirect way of representing this identity. This representation must instead depend on my apprehending a feature of my representations (or elements of them) (Allison 1983: 142–4; Guyer 1977: 267, 1987: 133–39; Patricia Kitcher 2011:147). The appropriate feature is a type of unity or ordering of these states. The idea is that if the representations I can attribute to myself possess a unity of the right kind, and if I apprehend or am cognitively sensitive to this unity, then I will be able to represent the apperceiving subject of any one of them as identical with that of any other.

Thus my representation of the identity of the subject comes about “only in so far as I conjoin one representation with another, and am conscious of the synthesis of them” (B133). This consciousness is profitably interpreted as conscious awareness not of the act or process of synthesis itself, but rather of the unity that is its outcome (Strawson 1966: 94–6; Dicker 2004: 133–4)). What sort of unity must I consciously recognize among my representations that would account for my representation of this identity? Note that it is not plausibly co-consciousness, for I represent the subject as identical for self-attributed representations that are not co-conscious, so actual co-consciousness could not explain generally how I represent this sort of identity. A credible alternative is that the unity consists in certain intimate ways in which representations in a single subject are typically related. Arguably, the essential feature of this unity is that a single subject's representations be inferentially and causally integrated to a high degree, and in this respect they are unified in a way in which representations possessed by discrete subjects are not. Alternatively, several commentators have argued that the relevant unity might be a temporal order among my representations, thereby linking the B-Deduction with the arguments of the Second Analogy and the Refutation of Idealism (Guyer 1977: 267; Dicker 2004: 137–44). A concern about this route is that a cognitive sensitivity to the time-ordering of representations does not obviously facilitate our representing them as belonging to a single subject (Brueckner 1984: 199–208). By contrast, when mental states fail to exhibit inferential and causal integration, as in the case of multiple-personality disorder, we have a tendency to posit multiple subjects, while we do not do so when such integration is present.

In Kant's view, the candidates for explaining how this kind of unity comes to be, or, less ambitiously,

for explaining my ability to recognize this sort of unity, are association and synthesis. At this point he appears to suppose that because Hume's psychological theory has already been ruled out, synthesis is the only remaining option. So for me to represent the identity of the subject of different self-attributions, I must generate or at least recognize the right sort of unity among these representations, and synthesis must be invoked to account for this unity. Thus Kant contends that this combination “is an affair of the understanding alone, which itself is nothing but the faculty of combining *a priori*” (B134–5. Since the understanding provides concepts for synthesis, and because for synthesis to be *a priori* is, at least in part, for it to employ *a priori* concepts, Kant is contending here that synthesis by means of *a priori* concepts is required to account for the unity in question.

Here is an austere representation of the structure of the argument so far:

1. I am conscious of the identity of myself as the subject of different self-attributions of mental states. (premise)
2. I am not directly conscious of the identity of this subject of different self-attributions of mental states. (premise)
3. If (1) and (2) are true, then this consciousness of identity is accounted for indirectly by my consciousness of a particular kind of unity of my mental states. (premise)
4. This consciousness of identity is accounted for indirectly by my consciousness of a particular kind of unity of my mental states. (1, 2, 3)
5. If (4) is true, then my mental states indeed have this particular kind of unity. (premise)
6. This particular kind of unity of my mental states cannot be accounted for by association. (5, premise)
7. If (6) is true, then this particular kind of unity of my mental states is accounted for by synthesis by *a priori* concepts. (premise)
8. This particular kind of unity of my mental states is accounted for by synthesis by *a priori* concepts. (6, 7)

Premise (1) is intended as a claim the skeptic about the legitimate applicability of *a priori* concepts will at least initially accept. The crucial necessary conditions, expressed by (3) and (7), are necessary conditions of only possible explanation. However, Paul Guyer forcefully argues that establishing the need for synthesis by means of *a priori* concepts would require ruling out the alternative explanation that empirical information and concepts derived from it is sufficient to account for the recognition of the unity at issue (Guyer 1987:146–7). And in his view it remains open, given what Kant has shown, that this recognition requires only awareness of information derived from inner sense or introspective experience. Kant does not attempt at this point in the argument to rule out such a rival empiricist hypothesis, but he arguably would need to do so to establish the need for *a priori* synthesis. To advance his claims, one might appeal to features of this unity that would render such an empiricist account inadequate. As we shall see, Kant employs this tactic in the next phase of the argument, which introduces his account of our representations of objects.

## 1.2 Representations of Objects and the Single-Substantive-Premise Interpretation of the B-Deduction

According to one widespread reading of the B-Deduction, §§15–20 comprise an argument whose only assumption is the premise about self-consciousness that Kant defends in §16. Strawson, for



example, is a proponent of such an interpretation (1966), as are Robert Paul Wolff (1963), Jonathan Bennett (1966), Henry Allison (1983), Edwin McCann (1985), and Dennis Schulting (2013). Demonstrating that we represent objects or an objective world has a key role in most versions of this reading. On Strawson's interpretation, for instance, “the fact that my experience is of a unified objective world is a necessary consequence of the fact that only under this condition could I be conscious of my diverse experiences as one and all my own” (1966: 94). Some interpreters dissent; Karl Ameriks, for example, contends that Strawson's conception of the Deduction is motivated by a desire to see it as showing that the skeptic about the external world is mistaken, while in fact refuting this sort of skeptic is not one of Kant's aims for this argument (Ameriks 1978, Allais 2011). Patricia Kitcher (2011: 115–18) argues against the single premise about self-consciousness interpretation on historical and textual grounds.

An uncontroversial role of §17 is to provide a characterization of an object, or more to the point, of a representation of an object, that facilitates a challenge to Humean associationism. Kant's proposal is that an object is “that in the concept of which a manifold of a given intuition is united” (B137). Here ‘object’ should be read in the broad sense of *objective feature of reality* — a feature whose existence and nature is independent of how it is perceived (B 142; Bird 1962/1973 130–31; Strawson 1966: 98–104; Guyer 1987: 11–24). Allison is a proponent of the view that §17 contain not only this challenge to Hume, but also an attempt to demonstrate that we represent objects on the basis of the conclusions about self-consciousness established in §16. This interpretation is a component of Allison's broader vision of the B-Deduction, according to which Kant demonstrates that the unity of apperception entails that we represent objects, and, conversely, that our representing objects entails the necessary unity of apperception (Allison 1983: 144ff.) Indeed, the crucial claim for Allison's interpretation is that the unity of apperception is not only a necessary but also sufficient condition for our representing of objects. This he calls the *reciprocity thesis*. Other commentators, including Richard Aquila (1989: 159), Howell (1992: 227–8) and Schulting (2013) agree that the B-Deduction features the reciprocity thesis and an attempt to establish its truth, while Ameriks (1978) disagrees. On an account of the sort Ameriks favors, the unity of apperception, and more exactly, the synthesis that explains our consciousness of the identity of the subject, is only a necessary condition for the representation of objects (cf. Allais 2011).

Allison's interpretation is attractive particularly because it promises leverage against the skeptic who denies that we represent objects, and also because this leverage is generated by premises about self-consciousness that this skeptic is likely to accept. Nevertheless, there are textual and charitable reasons to resist this reading (Ameriks 1978, Pereboom 1995; Patricia Kitcher 2011: 115–60). First of all, in §§18–20 Kant makes significant assumptions about features of our representations of objects that exceed anything that he has argued for in §17 or earlier. In particular, §18 he assumes that our representations of objects manifest a certain kind of necessity and universality, and this he does not purport to establish in §17 or earlier. Moreover, in the summary of the preceding steps of the B-Deduction in §20 Kant does not include premises from §§15–16. What we actually encounter in §20 indicates that Kant intends §§17–20, with some help from §13, to constitute a single, self-contained argument that does not depend on the conclusions about self-consciousness developed in §§15–16.

On Allison's conception, the argument from the unity of apperception (or, equivalently, from the unity of consciousness) for our representing objects occurs in the following passage:

(A) Understanding is, to use general terms, the faculty of cognitions (*Erkenntnisse*). They consist (*bestehen*) in the determinate relation of given representations to an object: and an object 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 cognitions (*Erkenntnisse*): and upon it therefore rests the very possibility of the understanding. (B137)

Allison himself presents a problem for his interpretation of this passage. He contends, first of all, that the reciprocity thesis is encapsulated in this sentence:

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

and that 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.

Allison points out that on this reading Kant's reasoning appears to involve a *non sequitur*, since (2) supports only Kant's having in mind that the unity of consciousness is a necessary condition for the representation of an object, and not for its also being a sufficient condition. Howell raises a similar concern: "In §17 Kant simply does not make this inference clear, and an air of blatant fallacy hovers over this part of his reasoning" (Howell 1992: 228; for an opposing view see Keller 1998: 80). But Ameriks argues that the B-Deduction should not be interpreted as providing an argument for the sufficiency claim, and a respectable case can be made for his reading (Ameriks 1978; Pereboom 1995; Patricia Kitcher 2011: 115-60).

Allison and Howell both argue that (1) should be read as a statement of the sufficiency claim. Now in (A) Kant contends that cognitions of objects consist in a determinate relation of representations to objects, and as (1) indicates, this relation is constituted or produced by a synthesis that essentially involves the unity of consciousness. However, (1) does not indicate that the synthesis that involves unity of consciousness cannot occur without its resulting in a relation of a representation to an object. By analogy, the smelting and molding of steel are processes that constitute or produce steel girders, but from this one should not conclude that the processes of smelting and molding steel cannot take place without the production of steel girders. Just as producing steel girders also requires molds of particular shapes, so producing representations of objects might require, in addition to the synthesis that involves the unity of consciousness, particular concepts of objects.

If the sufficiency claim, and with it the reciprocity thesis, is denied, we need an alternative account of how §17 functions in the overall argument of the B-Deduction. It may be that the role of this section is largely to provide a characterization of an object that has a key role in the ensuing challenge to Humean associationism, and thereby initiates an argument from below. Kant's proposal is that an object is "that in the concept of which a manifold of a given intuition is united" (B137). The subsequent claim is that the unification of a manifold requires synthesis; immediately following the characterization of an object he states that "all unification of representations demands unity of consciousness in the synthesis of them" (B137). It seems consistent with these texts that Kant's

characterization of an object is designed just to present his anti-Humean theory of the mental processing required for representing objects, and that the subsequent claim for the need for synthesis does not express a view he expects us to accept without further argument, but rather one he aims to confirm in §§18–20. If this argument succeeds, it will turn out that the *a priori* synthesis required to account for the features of our representations of objects Kant singles out is the same process that yields my consciousness of the identity of myself as subject of different self-attributions.

### 1.3 Universality and Necessity in our Experience of Objects

In §18 Kant draws our attention to certain features of our representations of objects that, in his view, will serve to defeat associationism and establish *a priori* synthesis:

The transcendental unity of apperception is that unity through which all the manifold given in an intuition is united in a concept of the object. It is therefore entitled objective, and must be distinguished from the subjective unity of consciousness... Whether I can become *empirically* conscious of the manifold as simultaneous or as successive depends on circumstances and empirical conditions. Therefore, the empirical unity of consciousness, through association of representations, itself concerns an appearance, and is wholly contingent... Only the original unity is objectively valid: the empirical unity of apperception,... which... is merely derived from the former under given conditions *in concreto*, has only subjective validity. One person connects the representation of a certain word with one thing, the other [person] with another thing; the unity of consciousness in that which is empirical is not, as regards what is given, necessarily and universally valid. (B139–40)

For Kant, a defining feature of our representations of objects is their objective validity. For a representation to be objectively valid it must be a representation of an objective feature of reality, that is, a feature whose existence and nature is independent of how it is perceived (Guyer 1987:11–24). In this argument, it appears that Kant just assumes that the representations that make up experience are objectively valid. He then aims to establish that association is inadequate because it can yield only representations that are not objectively valid. In the above passage, Kant contends that our objectively valid representations must in a sense be necessary and universal. However, the empirical unity of consciousness, which involves an ordering of representations achieved by association, can only be non-universal, contingent, and hence merely subjectively valid, by contrast with the transcendental unity of apperception, which involves an ordering that is universal and necessary, and is therefore objectively valid. In Kant's conception, it is the fact that the transcendental unity of apperception is generated by *a priori* synthesis that allows it to yield an ordering that is universal, necessary, and objectively valid. (Ameriks 1978, Pereboom 1995, Patricia Kitcher 2011:115-60; Allais 2011)

To illustrate and support these claims, Kant invokes examples of the ordering of phenomena in time that also have the key role in the discussion of the Second Analogy (cf. Guyer 1987: 87–90; Dicker 2004: 137–44). There Kant argues that our representations, considered independently of their content, are always successive. For example, when I view the front, sides, and back of the house when walking around it, and when I watch a boat float downstream, my representations of the individual parts and states occur successively. The objective phenomena represented by these successive representations, however, can be represented as either successive or as simultaneous — I represent the positions of the

boat as successive, but the parts of the house as simultaneous. More precisely, despite the representations in each of these sequences being *subjectively successive*, I represent the parts of the house as *objectively simultaneous*, and the positions of the boat as *objectively successive*. How do we account for this difference in objectivity despite the similarity in subjectivity? (Melnick 1973: 89)

The important clue for answering this question is that these representations of objective simultaneity and succession are universal and necessary. It is the universality and necessity of our representing the parts of the house as simultaneous that accounts for our representing them as objectively simultaneous, and the universality and necessity of our representing the positions of the boat as successive that accounts for our representing them as objectively successive. Association is inadequate for accounting for this objectivity because it is incapable of yielding such universality and necessity, a defect not shared by synthesis.

A first approximation of the import of ‘universal’ in the house example is:

(U) Any human experience of the parts of the house is an experience of these parts as objectively simultaneous.

The addition of necessity has the following effect on (U):

(U-N, first pass) Necessarily, any human experience of the parts of the house is an experience of these parts as objectively simultaneous.

This claim would be resisted by Hume if the necessity were specified as ranging over all possible circumstances, for Hume's theory would allow for the possibility of a deviant ordering in unusual empirical conditions. But (U-N, first pass) can be reformulated more precisely as

(U-N) Necessarily, if empirical conditions are normal, any human experience of the parts of the house is an experience of these parts as objectively simultaneous.

Kant's view is that given only the resources of association, the truth of (U-N) cannot be explained. The reason is “whether I can become empirically conscious of the manifold as simultaneous or as successive depends on circumstances or empirical conditions,” and so “the empirical unity of consciousness, through association of representations, itself concerns an appearance, and is wholly contingent” (B139–40). Association cannot explain the truth of (U-N), for given only the resources of association, the parts of the house will not necessarily or universally be represented as objectively simultaneous even supposing only normal empirical conditions. Kant asks us to consider an activity, word association, which functions as a paradigm for association. Word association, familiarly, does not yield universal and necessary patterns; “one person connects the representation of a certain word with one thing, the other [person] with another thing...” (B140). Hume's own paradigm for association is in the relations among parts of a conversation (1748: §3). In such conversations, people make different associations in the same circumstances. Kant's point is that if the very paradigms for association fail to exhibit the sort of necessity and universality at issue, then the hypothesis that association is powerful enough to yield such an ordering of representations — wherever we find it — is ruled out. (Pereboom 1995; Dickerson 2004: 170–77; for an account of a more general relationship in Kant between universality and necessity on the one hand and apriority on the other, see Smit 2009).

Here we should see Kant as advancing his claim for the applicability of the categories by ruling out association as an explanation for (U-N). The structure of this (part of the) argument can be represented as follows:

9. We have representations of objects, i.e., of objectively valid phenomena. (premise)
10. All of our representations of objects are (in part) of universal and necessary features of experience. (9, premise)
11. Necessary and universal features of experience cannot be explained by association. (premise, from reflection on the nature of association))
12. If (10) and (11) are true, all of our representations of objects require a faculty for ordering mental states distinct from association. (premise)
13. All of our representations of objects require a faculty for ordering mental states distinct from association. (11, 12)
14. If (13) is true, all of our representations of objects require a faculty for synthesis by *a priori* concepts. (premise)
15. All of our representations of objects require a faculty for synthesis by *a priori* concepts. (13, 14)

We can expand (15) to explicitly note the link to the argument from above:

- 15\*. All of our representations of objects require a faculty for synthesis by *a priori* concepts, the same faculty required to account for my consciousness of the identity of myself as subject of different self-attributions of mental states. (1-8, 13, 14)

To this we can add the final moves, which are explained in the subsequent sections of the B-Deduction:

16. Insofar as our representations of objects require a faculty for synthesis by *a priori* concepts, certain *a priori* concepts -- the categories -- legitimately apply to these objects. (premise)
- C. We have representations of objects, and they are all such that the categories legitimately apply to these objects. (9, 15, 16)

The key necessary conditions, expressed by (12) and (14), like those of the argument from above, are necessary conditions of only possible explanation.

Guyer objects that at various places in the Transcendental Deduction Kant illegitimately assumes knowledge of necessity, and perhaps this argument falls to such a concern (Guyer 1987: 146–7). While this concern has the potential of weakening Kant's argument, perhaps Hume would not deny the necessity under consideration at this point in the argument. Hume does maintain that it is in some sense impossible, given an experience of constant conjunction, that the mind not be carried from an impression of the first conjunct to an idea of the next:

... having found, in many instances, that any two kinds of objects, flame and heat, snow and cold, have always been conjoined together; if flame or snow be presented anew to the senses, the mind is carried by custom to expect heat or cold, and to *believe*, that such a quality does exist, and will discover itself upon a nearer approach. This belief is the

necessary result of placing the mind in such circumstances. It is an operation of the soul, when we are so situated, as unavoidable as to feel the passion of love, when we receive benefits; or hatred, when we meet with injuries. (1748: §5)

Thus Hume himself contends that given certain specific empirical circumstances, a particular type of ordering of perceptions in a sense necessarily comes about.

A further concern of Guyer's is that Kant assumes without defense that all knowledge of necessity is grounded in *a priori* concepts. But, in response, perhaps we need not interpret Kant as arguing directly from (U-N) to the claim that the categories correctly apply to objects in our experience. Rather, we might see him as advancing his claim for the applicability of the categories by ruling out association as an explanation for (U-N). This transition can be divided into three steps:

- a. To explain the truth of (U-N), we must have a mental faculty other than association for ordering representations.
- b. This faculty does not consist solely of sensory items.
- c. This faculty must employ *a priori* concepts, the categories in particular.

The challenge Kant issues is to explain why, under normal conditions, the ordering in question is universal and necessary. Part of the best explanation, he believes, is (a), that we must have a faculty for ordering the representations. Hume might agree with this conclusion, supposing a sufficiently thin conception of 'faculty'. Yet he would deny (b), that this faculty does not consist solely of sensory items. Kant argues that the Humean proposal for a faculty that consists solely of sensory items, the faculty of association, cannot account for the truth of propositions such as (U-N), for the very paradigms of association, such as word association, and the association of topics in a conversation, do not exhibit the requisite universality and necessity. The alternative that can account for the truth of propositions such as (U-N) involves affirming (c), that the faculty in question must be one that employs the categories.

The associationist might counter that sensory experience is sufficiently uniform for association to produce the universalities and necessities at issue. Perhaps Kant is too quick to conclude that the argument from universality and necessity is decisive, for in addition, associationist objections of this sort must be answered — as contemporary critical discussions of proposals for innate concepts indicate (Pereboom 1995: 31–3). But this does not detract from the anti-associationist force provided by the sorts of universalities and necessities Kant has in mind, and this fact is recognized by the contemporary discussion.

## 1.4 Logical Forms of Judgment and the Categories

In §§19–20, Kant contends that the vehicle that brings about the synthesis at issue is judgment, and that this vehicle employs certain forms of judgment, which are in turn intimately related to the twelve categories. By connecting synthesis to judgment in this way, and the forms of judgment to the categories, Kant aims to show that we must employ the categories in the synthesis of our experience of objects.

In §19, Kant argues that there must be a certain way in which each of my representations is unified in

the subject, and he identifies this way with judgment: “I find that a judgment is nothing but the manner in which given cognitions are brought to the objective unity of apperception” (B141). Judgment, Kant proposes, is objectively rather than subjectively valid, and hence exhibits the type of universality and necessity that characterizes objective validity (B142). He then claims that without synthesis and judgment as its vehicle, an ordering of representations might reflect what appears to be the case, but it would not explain how we make distinctions between objective valid phenomena (i.e., objects) and the subjective states they induce.

In §20, Kant ties this notion of judgment to the twelve forms of judgment presented in the Metaphysical Deduction (A70/B95), and he then connects these forms of judgment to the twelve categories (A76–83/B102–9). The claim has often been made that the links Kant specifies between synthesis and judgment, judgment and the forms of judgment, the forms of judgment and the categories are not sufficiently supported. Guyer, for example, argues that Kant has not adequately established the last of these connections, that although Kant claims that the categories are simply the forms of judgment as they are employed in the synthesis of representations in an intuition (A79/B104–A80/B105, B143), he has failed to make this claim plausible (Guyer 1987: 94–102). It is fair to say that these concerns have merit. Kant's assertions about these ties remain more obscure than the preceding part of the Transcendental Deduction, and it continues to be a serious challenge for interpreters to clarify and vindicate them.

Béatrice Longuenesse (1998), in her interpretation of the Metaphysical Deduction, takes up this challenge. In her view, the faculty at issue in the production and use of concepts, the understanding, is the power to judge (*Vermögen zu Urteilen*), which is ultimately a disposition or a *conatus* to make judgments and to shape how we are affected so that we can make them (Longuenesse 1998: 208, 394). The logical forms of judgment are in essence the forms of combination of concepts in judgments. One such form is the categorical, which is the form of subject-predicate judgments; another is the hypothetical, the form of conditional judgments. Kant contends that the logical form of a judgment is what makes it capable of truth or falsity, for by means of such a form a judgment can constitute a relation of a subject's representations to an objective feature of reality (Longuenesse 2000, 93–4). For instance, by its categorical form, the judgment ‘the boat is moving’ can constitute the relation of my representations of a boat and of motion to an objectively existing boat in motion, and as a result this judgment can be true or false.

One role of the logical forms of judgment is in the process of analysis, by which the objects we intuit are subsumed under concepts. What results from this process is a judgment that expresses what Kant calls an *analytic unity* — paradigmatically, the unity in the subsumption of several intuited objects under a single concept. But a logical form of judgment can also function in a different role: in the synthesis of a manifold of an intuition. The result in this case is a *synthetic unity*, the unity of a synthesized multiplicity of representations in a single intuition. The understanding, as the power to judge, functions in each of these two roles; “the same function that gives unity to concepts in judgment, also gives unity to the mere synthesis of representations in intuition” (A79/B104–5). In its synthetic role, the understanding adds content to the forms of judgment:

The same understanding, and indeed by means of the very same actions through which it brings the logical form of judgment into concepts by means of the analytical unity, also brings a transcendental content into its representations by means of the synthetic unity of

the manifold in intuition in general. (A79/B105)

The addition of such transcendental content turns the form of judgment into a category. This content is a feature of the forms of intuition, space and time, which are called upon when the power to judge sets out to unify a manifold of intuition (B128–9). The categories — more precisely, the versions of the categories that are schematized for our way of cognizing (A137/B176ff) — are thus generated from the forms of judgment in the process of synthesizing intuitions by the addition of spatial and temporal content. For example, generated from the categorical form of judgment by the addition of such content is the category of substance, and generated from the hypothetical form of judgment in this way is the category of cause. (For beings possessing the power of understanding but with different forms of intuition, the categories would be schematized differently from ours.)

## 1.5 The Second Step of the B-Deduction

In §20 Kant, draws a conclusion from the considerations he has so far advanced: “Consequently, the manifold in a given intuition is necessarily subject to the categories” (B143). One might think that this is precisely what Kant intended to establish in the Transcendental Deduction, and thus that the argument is brought to an end in §20. However, in §21 he indicates that the Deduction is not yet complete: “Thus in the above proposition a beginning is made of a deduction of the pure concepts of understanding” (B144). Kant goes on to explain:

In what follows (cf. §26) it will be shown, from the mode in which the empirical intuition is given in sensibility, that its unity is no other than that which the category (according to §20) prescribes to the manifold of a given intuition in general. Only thus, by demonstration of the *a priori* validity of the categories in respect of all objects of our senses, will the purpose of the deduction be fully attained. (B144–5)

Here a perennial interpretive question arises: how should we construe the argument we find in §26, together with material from §24, which is sometimes designated as the second step of the B-Deduction?

One position on this interpretive issue is advanced by Erich Adickes (1889: 139–4) and H. G. Paton (1936: v. 1, 501), who argue that while the material that precedes §21 constitutes an *objective* deduction, material in §24 and §26 comprise a *subjective* deduction. This distinction has its source in the Preface to the A edition:

This enquiry, which is somewhat deeply grounded, has two sides. The one refers to the objects of pure understanding, and is intended to expound and render intelligible the objective validity of its *a priori* concepts. It is therefore essential to my purposes. The other seeks to investigate the pure understanding itself, its possibility and the cognitive faculties upon which it rests: and so deals with it in its subjective aspect. Although this latter exposition is of great importance for my chief purpose, it does not form an essential part of it. (Axvi–xvii)

Henrich (1968–69) rejects the Adickes/Paton proposal for the reason that in §21 Kant states that the demonstration of the validity of the categories is completed only in §26, and the passage from the



A-Preface indicates that this is a task for the objective deduction. In defense of Adickes and Paton, in §20 Kant claims to have established that the categories apply to the manifold in any one given intuition, and affirms that he will now show that categories apply to any object presented to the senses. In view of the aim of the objective deduction, this move would seem to require only a straightforward and trivial application of the result of §20 to any empirical intuition we might have (Dickerson 2004: 195–96). So, on a charitable interpretation of Kant's agenda, showing that the categories apply to any object presented to the senses plausibly involves considerations beyond the scope of the objective deduction.

Henrich points out that although in the B-Deduction Kant sought to avoid the problems of a subjective deduction, this “does not mean that he neglected the demand for an explanation of the possibility of relating the categories to intuitions.” However, by Kant's account, a subjective deduction features not only an examination of cognitive faculties, but also an investigation of “the possibility of the pure understanding,” which would include an investigation of how the categories are related to empirical intuition and to the objects presented to us in such intuition. This is precisely the type of inquiry that we encounter in §24 and §26 — and Henrich agrees. Thus, in the last analysis, it seems that the views of Henrich, Paton, and Adickes can be reconciled: they concur that the second step aims to show how it is that the categories are related to objects of experience -- in such a way as to show how the categories correctly apply to them. Moreover, this common ground yields a satisfying interpretation of the specifics of the text. In §26 Kant contends that our intuitive representations of space and time, since they themselves contain a manifold, must also be synthesized by means of the categories. And now, because all objects that can be presented to us in intuition are given to us in space and time, our representations of these objects will also be synthesized by the categories. Kant thus provides an explanation of how the categories apply to all the objects that can be presented to the senses, or in experience, by way of the manner in which these objects are given to us (cf. Longuenesse 1998: 211–33; Keller 1998: 88–94; Dickerson 2004: 196–201; Pollok 2008).

The text of the second step gives rise to a number of general important interpretive issues for Kant's overall position. One such issue, raised by the complex footnote at B160–61, is whether Kant countenances a kind of unity in our representations which is not a result of synthesis by *a priori* concepts, but is instead an upshot of the forms of intuition (McLear 2014) or of the productive imagination independently of the categories (Heidegger 1929/1997; cf., Longuenesse 1998). What one concludes about the footnote has implications for whether Kant is a conceptualist or a nonconceptualist about intuitions or our representations of objects (Hanna 2008, 2011, Grüne 2011, Schulting 2012, McLear 2014). Robert Hanna (2011) argues that the possibility of nonconceptual intuition results in a gap in the argument of the B-Deduction, since this would allow for intuitions to which the categories do not apply. Stefanie Grüne (2011) and Schulting (2012) oppose this line of reasoning.

## 2. The Refutation of Idealism

In the transcendental argument of the Refutation of Idealism, Kant's target is not Humean skepticism about the applicability of *a priori* concepts, but rather Cartesian skepticism about the external world. This argument was added by Kant to the second edition of the *Critique of Pure Reason* (B274–279, with a change suggested in the Preface to B, Bxxxix–Bxli), and subsequently embellished and

reworked in a series of Reflections (Guyer 1987: 279–316). More specifically, Kant intends to refute what he calls problematic idealism, according to which the existence of objects outside us in space is “doubtful and indemonstrable” (B274). His strategy is to derive the claim that such objects exist from my awareness that my representations have a specific temporal order. At the present time I am aware of the specific temporal order of many of my past experiences, an awareness produced by memory. But what is it about what I remember that allows me to determine the temporal order of my experiences? There must be something by reference to which I can correlate the remembered experiences that allows me to determine their temporal order. But first, I have no conscious states that can play this role. Further, this reference cannot be time itself, for “time by itself is not perceived;” Kant argues for this claim in the First Analogy (B224–5). As Guyer puts it, it is not as if the content of memories of individual events are manifestly indexed to specific times, as sportscasts and videotapes often are (Guyer 1987: 244). Kant contends that the only other candidate for this reference is something outside of me in space, and it must be something (relatively) permanent (cf. First Analogy, B224–5). Perhaps this claim is made plausible by how we often actually determine the times at which our experiences occur. We use observations of sun's positions, or of a clock that indicates time by way of the period of a pendulum, or by the period of the vibration of a cesium atom. Kant's argument can be seen as exploiting this fact, together with the observation that there is no similar periodic process in human conscious experience considered independently of any spatial objects it might represent, and that we lack any awareness of time by itself, to show we must perceive objects outside us in space by reference to which we can determine the temporal order of our past experiences.

George Dicker provides a compelling initial representation of Kant's argument (Dicker 2004, 2008):

1. I am conscious of my own existence in time; that is, I am aware, and can be aware, that I have experiences that occur in a specific temporal order. (premise)
2. I can be aware of having experiences that occur in a specific temporal order only if I perceive something permanent by reference to which I can determine their temporal order. (premise)
3. No conscious state of my own can serve as the permanent entity by reference to which I can determine the temporal order of my experiences. (premise)
4. Time itself cannot serve as this permanent entity by reference to which I can determine the temporal order of my experiences. (premise)
5. If (2), (3), and (4), are true, then I can be aware of having experiences that occur in a specific temporal order only if I perceive persisting objects in space outside me by reference to which I can determine the temporal order of my experiences. (premise)
6. Therefore, I perceive persisting objects in space outside me by reference to which I can determine the temporal order of my experiences. (1–5)

Several points of interpretation should be noted. First, it is generally agreed that the notion of awareness in Premise (1) should be interpreted as a success notion, i.e., that to be aware that I have experiences that occur in a specific temporal order is to correctly determine that they occur in this order (Allison 1983: 290ff, Guyer 1987: 293ff, Dicker 2004, 195ff., 2008; Chignell 2010). Second, most commentators concur that the experiences at issue are my past experiences. Guyer suggests that the argument might be extended to combinations of my past and my present experience, but Dicker objects that the fact that past experiences occurred before the present one is directly knowable by introspection, and so doesn't require the external reference (Guyer 1987; Dicker 2004: 195ff.). Third, Jonathan Bennett points out that we have single memories whose content spans an appreciable length

of time that allow us to determine the order of past mental states. When one remembers hearing a certain word, one can accurately determine by the content of a single memory that certain phoneme-experiences occurred prior to other phoneme-experiences — for example, that one heard ‘mad’ and not ‘dam’ (Bennett 1966: 228–9). Dicker (2004: 201–2) remarks that we may be able to determine accurately the order some of our past states by means of such memories, but we cannot ascertain the order of most of our past experiences in this way. He therefore advocates a restriction of the argument to the experiences we can correctly order but not in the way Bennett adduces. Fourth, many commentators have noted that we might have perceived time directly. Guyer, as we just saw, suggests that all of our conscious experiences might have featured a time clock, much like a television sportscast or a video camera (1987: 244; cf. Strawson 1966). But as Dicker points out, in actual fact our experience does not have any such feature, and he is content for Premise (4), that time itself cannot serve as the reference whereby I correctly determine the temporal order of my past experiences, to state a merely contingent fact about us (2004, 2008). Chignell (2010) expresses a concern about interpretations of the Refutation in which it is merely a contingent fact about us that the alternative methods for determining the temporal order of my past experiences are unavailable, for then the conclusion, that we perceive objects in space, would inherit such mere contingency, which he argues to be at odds with Kant's hopes for the Refutation. Fifth, commentators differ on the relation that must obtain between the objects in space and the experiences whose temporal order we can correctly determine. Guyer argues that the relation must be causal, since “the states of the self are judged to have a unique order just insofar as they are judged to be caused... by the successive states of enduring objects” (Guyer 1987: 309; cf. Dicker 2004: 200, 2008). By contrast, in Allison's view we require the temporal order of objects in space only as a backdrop against which to determine the temporal order of our experiences; “an enduring, perceivable object (or objects) is required to provide a frame of reference by means of which the succession, coexistence, and duration of appearances in a common time can be determined” (Allison 1983: 201). If the object in space that provides the reference is the sun, for example, the states of the sun don't need to cause my experiences for me to determine their temporal order by means of those states.

Three of the most pressing problems that have been raised for the Refutation are the following. First, a skeptic might well reject Premise (1) on the ground of a general skepticism about memory (Allison 1983: 306–7). Bertrand Russell, for example, suggests that for all I know I was born five minutes ago (Russell 1912). On this skeptical hypothesis, I would be mistaken in my belief that I had experiences A, B, and C which occurred more than five minutes ago, first A, then B, and lastly C. Plausibly, a skeptic who claimed that we lack adequate justification for a belief that external objects exist would likely also be disposed to claim that I lack justification for my belief that I had experiences that occurred in the past in that specific temporal order. Thus Kant's supposition that Premise (1) yields leverage against an external-world skeptic is mistaken (cf. Dicker 2008, Chignell 2010).

Second, consider the claim that my mental states (or the mental subject itself) are as suited as objects in space to function as a reference whereby I can correctly judge the temporal order of my past experiences. Suppose I had available as such a reference only the mere mental appearance of a digital clock in one corner of my field of consciousness. This would not clearly be less effective than an actual clock in space. To this one might reply, with Dicker, that there are no actual mental states that are adequate to serve as such a reference. However, and this is the deeper worry, on Berkeley's idealist view according to which the *esse* (to be) of objects in space is their *percipi* (to be perceived), any spatial objects would be nothing more than mental states of some subject, or aspects of those states,

but Berkeleyan spatial perceptions would be as adequate a reference by which to determine the temporal order of my past experiences as perceptions of spatial objects that are in some sense distinct from my mental states (cf. Allison 1983: 300–1; Vogel 1993; Chignell 2010).

Third, one might contend that Kant's Refutation demonstrates that the reference in question must be (relatively) permanent, and that there is nothing in the Berkeleyan spatial realm that satisfies this requirement. However, to this one might respond that the reference by which I determine the temporal order of my past experiences need not be permanent in a way that cannot be satisfied by Berkeleyan spatial objects. If in the corner of my field of consciousness featured distinct momentary flashes, every second, indicating the date and time to the second, I would be able to determine the temporal order of my past experiences by their means (cf., van Cleve, reported in Dicker 2004: 207).

A note on the second of these concerns: Several commentators have argued that Kant's Refutation of Idealism is meant to undermine any metaphysical sort of idealism, including metaphysical idealisms that have been attributed to Kant (Guyer 1987: 317–29). But while this reading has interesting support, one should hesitate to endorse it solely on the ground that Kant maintains that spatial objects are distinct from the states of the perceiver. On a plausible metaphysical interpretation of Kant's idealism, the *esse* of objects in space is not their *percipi*, since Kant's spatial objects, by contrast with Berkeley's, are recognition-transcendent. In Berkeley's position, a subject's perception of an oar in the water as crooked is not a misperception, for “what he immediately perceives by sight is not in error, and so far he is in the right,” and it is misleading only because it is apt to give rise to mistaken inferences (Berkeley 1713: Third Dialogue); while for Kant this perception is in error. In Kant's view, the oar is recognition-transcendent by virtue of the Second Postulate's provision that the actuality of such objects is determined causally:

That which is bound up with the material conditions of experience, that is, with sensation, is actual (*wirklich*). (A218=B266)

The actual is that which conforms to the system of empirical causal laws (A225/B272ff.), and because the crookedness of the oar does not so conform, it is not actual. How the Second Postulate rules out the existence of the apparent spatial objects of dreams and hallucinations is spelled out in the third note to the Refutation of Idealism:

Note 3. From the fact that the existence of outer things is required for the possibility of a determinate consciousness of the self, it does not follow that every intuitive representation of outer things involves the existence of those things, for their representation can very well be the product merely of the imagination (as in dreams and delusions). Such representation is merely the product of previous outer perceptions, which, as has been shown, are possible only through the reality of outer objects... Whether this or that supposed experience be not purely imaginary, must be ascertained from its special determinations, and through its congruence with the criteria of all actual experience (*wirklichen Erfahrung*). (B278–9 cf. A376, cf. A492=B520–1)

The objects of dreams and hallucinations don't meet the criterion of actuality of the Second Postulate, that is, according with the laws of causality. Thus while in Berkeley's view the *esse* of spatial things consists in their being perceived by the senses, for Kant they are the objects of the correct

causal/scientific account of the contents of our outer experience. This causal criterion allows that a spatial object, or, more precisely, the matter from which it is constituted, is (at least relatively) permanent — which would seem ruled out in Berkeley's position, since the spatial object does not last any longer than the idea does. (Berkeley does say that the ideas of what “are called real things” differ from “those excited in the imagination” by “being less regular, vivid, and constant” (Berkeley 1710, *Principles* Part I, 33), but for him this difference does not challenge or undermine the claim that the *esse* of a physical object is its *percipi*.) None of this forces a metaphysical realist as opposed to an idealist reading of Kant.

Yet at the same time, according to the second concern, the Refutation is inadequate even if its ambition is restricted to demonstrating the existence of merely metaphysically ideal but nonetheless recognition-transcendent objects outside us in space. For granting that the Refutation establishes that for me to determine the temporal order of my past experiences I must perceive objects in space, it fails to show that I need to perceive spatial objects any more realistic than the Berkeleyan ones. More generally, the worry is that Berkeleyan version of idealism has the resources to yield as adequate a reference for me to determine the temporal order of my past experiences as any rival position Kant is plausibly interpreted as endorsing, and that for this reason the Refutation falls short of its aim.

### 3. Contemporary Kantian Transcendental Arguments

Kant-inspired transcendental arguments against skepticism about the external world were developed with vigor in the mid-twentieth century, notably by P. F. Strawson, most famously in his Kantian reflections in *The Bounds of Sense* (1966). These arguments are often reinterpreted as, or at least inspired by, Kant's Transcendental Deduction and his Refutation of Idealism. Some are more ambitious than Kant's would seem to be, insofar as they attempt to refute some variety of skepticism by showing that there is an essential commitment of the skeptical position for which the falsity of that position is a necessary condition (Nagel 1997: 60ff.). Strawson's most famous transcendental argument (1966: 97–104) is modeled on the Transcendental Deduction, but explicitly without a commitment to synthesis or any other aspect of Kant's transcendental psychology. His target is a purely sense-datum experience, which does not feature objects, “conceived of as distinct from any particular states of awareness of them,” that is, a Berkeleyan experience of spatial objects whose *esse* is *percipi* (1966: 98). The core of the argument is as follows. From the premise that every (human) experience is such that it is possible for its subject to become aware of it and ascribe it to herself, we can infer that in every experience the subject must be capable of distinguishing a recognitional component not wholly absorbed by, and thus distinct from, the item recognized (1966: 100). From this we can infer that the subject must conceptualize her experiences in such a way so as to contain the basis for a distinction between a subjective component — how the experienced item seems to the subject, and an objective component — how the item actually is. Strawson specifies that “collectively,” this comes to “the distinction between the subjective order and arrangement of a series of such experiences on the one hand, and the objective order and arrangement of the items of which they are the experiences on the other” (1966: 101). Conceptualizing experience as involving an objective order and arrangement of items amounts to making objectively valid judgments about it, which, in turn, implies that experience must consist of a rule-governed connectedness of representations (1966: 98). Summarizing, from a premise about self-consciousness, we can infer that the subject must conceptualize her experience so as to feature a distinction between “the subjective

route of his experience and the objective world through which it is a route,” where the experience of the objective world consists in a rule-governed order of representations (1966: 105).

Barry Stroud, in his 1968 article “Transcendental Arguments,” issued a formidable challenge to the enterprise of defeating the external-world skeptic by transcendental arguments of this sort (Stroud 1968). There Stroud contends that such transcendental arguments are undermined by a problem which can be stated quite generally. These arguments feature reasoning from some aspect of experience or knowledge to the claim that the contested feature of the external world in fact exists. In each case the *existence* of the external feature will not be a necessary condition of the aspect of experience or knowledge in question, for a *belief* about the external feature would always suffice. Although the claim about existence of the aspect of the external world could be secured if certain kinds of verificationism or idealism were presupposed, these views are highly controversial. Besides, one could make as much of an inroad against the skeptic armed with the verificationism or idealism alone, without adducing the transcendental argument at all (cf. Brueckner 1983, 1984).

Although Strawson's transcendental argument in *The Bounds of Sense* is not a specific target in Stroud's (1968), Anthony Brueckner suggests that it is susceptible to the line of criticism that Stroud develops (Brueckner 1983: 557–8). For, arguably, Strawson can only conclude that experience must be conceptualized in a certain way, such as to allow the subject to make the distinction between an objective world and her subjective path through it. This is not a conclusion about how a mind-independent world must be, but only about how it must be thought. (Alternatively, Strawson might be read as drawing only a conclusion about how experience must be conceptualized, which would render the transcendental argument as one of a more modest variety (see below)).

Recent developments of transcendental arguments reflect a struggle to accommodate Stroud's criticism, and often chastened expectations about what such arguments might establish by way of refuting skepticism. One more modest sort of transcendental argument begins with a premise about experience or knowledge that is acceptable to the skeptic in question, and then proceeds not to the existence of some aspect of the external world, but in accord with Stroud's criticism, to a belief in the existence of some aspect of the external world. Stroud himself advocates a strategy of this sort (Stroud 1994, 1999), as does Robert Stern (1999b). The kind proposed by Stroud begins with the premise that we think of the world as being independent of us, and it concludes, as a necessary condition of this premise, that we must think of it as containing enduring particulars. Such an argument does not claim that as a necessary condition of this premise there must exist such particulars. It contends only for “a connection solely within our thought: if we think in certain ways, we must think in certain other ways” (Stern 1999b: 165). A belief or thought to which one reasons in this way would, in Stroud's conception, have a certain *indispensability*, “because no belief that must be present in any conception or any set of beliefs about an independent world could be abandoned consistently with our conception of the world at all,” and it would be *invulnerable* “in the special sense that it could not be found to be false consistently with its being found to be held by people” (Stern 1999b: 166).

Stern advances a conception of this kind of argument on which it addresses a skeptic who questions whether certain beliefs cohere with others in one's set, as opposed to a skeptic who questions whether certain beliefs are true (Stern 1999b). A modest transcendental argument would then aim to show that a belief whose coherence with the other beliefs is challenged so coheres after all. The requisite coherence might be demonstrated by showing that the belief in question is actually a necessary

condition of a belief that is indispensable (in some coherentist sense) to one's set. Mark Sacks (1999) objects that if at the same time one were to admit that the belief might not be true, one's sense that one is justified in holding the belief would be undermined. This worry is a serious one. Sacks argues that it arises because of a tension between the coherentist theory of justification and the realist correspondence theory of truth that the external world skeptic presupposes. One might respond, he points out, by accepting a coherence theory of truth as well, but this would be to adopt a version of idealism. Moreover, it seems that Sacks's proposal would not solve the problem, for the reason that even if one accepted a coherence theory of truth, one would still have to admit that for specific instances of a belief one might be mistaken, even if one maintained that one was justified in holding that belief on grounds of coherence.

An important criticism of Stroud's proposed sort of modest transcendental argument is raised by Brueckner (1996). Brueckner addresses the fit between the claim that certain non-skeptical beliefs are invulnerable in Stroud's sense, and the admission that they might not be true. More precisely, he challenges the claim that one can simultaneously affirm the following two principles, each gleaned from Stroud (1994):

(CT) If we attribute beliefs to speakers (if we believe that they have beliefs with determinate contents), then we must also believe that there is an independent world of enduring objects with which they interact.

and

(SK) Although we believe many things about a world independent of us and our experiences, ... none of those beliefs is true.

Brueckner divides (SK) into

i. we believe many things about an external reality independent of us and our experiences

and

ii. none of these beliefs is true.

Stroud argues, in effect, that given CT we cannot believe that (i) and (ii) are true, but nonetheless, we can believe that (i) and (ii) are logically compossible given CT. Brueckner argues that using similar reasoning, given CT we cannot conceive of a possible world in which both (i) and (ii) hold, and this fact undermines Stroud's claim. To conceive of a world W in which (i) is true is to conceive of a world in which we (in the actual world) attribute beliefs about mind-independent objects to counterfactual versions of us (CVs) in W. But given CT, we must now also conceive of W as featuring mind-independent objects with which these CVs interact. Consider the CVs' belief *that there exist mind-independent objects* — a belief they share with us. This belief of theirs will be true. And thus in W (ii) will be false. Hence we have not conceived of a world in which given CT, SK is nevertheless true, and indeed, we will not be able to conceive of such a world. But our not being able to conceive of a possible world in which given CT, SK is true, constitutes strong evidence against the claim that SK is logically possible given CT (1996: 274–75). Brueckner does not think that this argument demonstrates that it is inconsistent to accept CT and assert that SK is logically possible, but that there

will be no evidence of the usual kind for this claim — evidence from conceivability. Thus given the transcendental argument he advances, Stroud will be pushed in the direction of the immodest conclusion that it is not possible for us to be mistaken in our belief that there exist mind-independent objects.

From this one might be tempted to conclude that despite his critique of 1968, Stroud, with Brueckner's assistance, has found a transcendental argument that does in fact establish a conclusion about the external world. However, so far nothing has been said to turn back Stroud's 1968 critique, and it seems much more likely that the argument Stroud now advances can at best conclude with a version of CT in which the term “independent” must be read in a transcendently ideal sense — in which, for example, the nature of certain physical objects is determined by our best scientific theories about them, and our sensory experiences can be in error about these objects. And if this is so, then the argument would secure a belief about the external world only on the presupposition of metaphysical idealism, and this is one of the ways in which anti-skeptical transcendental arguments might be doomed according to Stroud's 1968 critique.

Despite these sorts of challenges, the aspiration to forge transcendental arguments with considerable anti-skeptical force has not waned. Qassim Cassam (1999), Sacks (2000), and Stern (2000), for example, have developed creative and nuanced versions of transcendental arguments designed to negotiate the type of problem Stroud has pressed.

## 4. Concluding Remarks

The legacy of the arguments such as the Transcendental Deduction and the Refutation of Idealism includes not only Kant's actual successes, but also a number of influential philosophical strategies: the now-standard tactic of arguing for concepts whose source is in the mind from universal and necessary features of experience; the idea of drawing significant philosophical conclusions from premises about self-consciousness alone; and the notion of a transcendental argument, which from an uncontroversial premise about our thought, knowledge or experience, reasons to a substantive and unobvious presupposition and necessary condition of this claim.

## Bibliography

### Primary Literature: Works by Kant

- Gesammelte Schriften*, ed. *Königlichen Preussischen Academie der Wissenschaften*, 29 Volumes, Berlin: Walter de Gruyter *et al.*, 1902–
- 1781/1787, *Critique of Pure Reason* (trans. P. Guyer and A. Wood). Cambridge and New York: Cambridge University Press, 1987. (References to *CPR* are in the standard pagination of the 1<sup>st</sup> (A) and 2<sup>nd</sup> (B) editions. A reference to only one edition means that the passage appeared only in that edition.)
- 1788, *Critique of Practical Reason* (trans. M. Gregor). Cambridge and New York: Cambridge University Press, 1997.
- 1790, *Critique of the Power of Judgment* (trans. P. Guyer and E. Matthews). Cambridge and New York: Cambridge University Press, 2001.



1804, *Opus Postumum* (trans. E. Förster and M. Rosen). Cambridge and New York: Cambridge University Press, 1995.

## Secondary Literature





- Adickes, E., 1889. *Immanuel Kants Kritik der reinen Vernunft*, Berlin: Mayer & Müller.
- Allais, L., 2011. "Transcendental Idealism and the Transcendental Deduction," in *Kant's Idealism: New Interpretations of a Controversial Doctrine*, D. Schulting and J. Verburgt, (eds.), Dordrecht: Springer, pp. 91–107.
- Allison, H., 1983. *Kant's Transcendental Idealism*, New Haven: Yale University Press; second edition, 2004.
- Aquila, R., 1989. *Matter and Mind*, Bloomington: Indiana University Press.
- Ameriks, K., 1978. "Kant's Transcendental Deduction as a Regressive Argument," *Kant-Studien*, 69: 273–87.
- Baum, M., 1986. *Deduktion und Beweis in Kants Transzendentalphilosophie*, Königstein, Athenaeum.
- Bennett, J., 1966. *Kant's Analytic*, Cambridge: Cambridge University Press.
- Bennett, J., 1979. "Analytic Transcendental Arguments," in Bieri, Horstmann and Krüger (eds.) 1979, pp. 45–64.
- Berkeley, G., 1710. *A Treatise Concerning the Principles of Human Knowledge*, Jonathan Dancy (ed.), Oxford: Oxford University Press, 1998.
- Berkeley, G., 1713. *Three Dialogues between Hylas and Philonous*, Jonathan Dancy (ed.), Oxford: Oxford University Press, 1998.
- Bieri, P., Horstmann, R., and Krüger, L. (eds.), 1979. *Transcendental Arguments and Science: Essays in Epistemology*, Dordrecht: D. Reidel.
- Bird, G., 1962/1973. *Kant's Theory of Knowledge*, London: Routledge and Kegan Paul; reprinted, New York: Humanities Press, 1973.
- Brueckner, A., 1983. "Transcendental Arguments I," *Noûs*, 17: 551–75.
- Brueckner, A., 1984. "Transcendental Arguments II," *Noûs*, 18: 197–225.
- Brueckner, A., 1996. "Modest Transcendental Arguments," *Philosophical Perspectives*, 10: 265–80.
- Caird, E., 1889. *The Critical Philosophy of Kant*, New York: Macmillan.
- Carl, W., 1989. "Kant's First Drafts of the Deduction of the Categories," in Förster 1989, pp. 3–26.
- Carl, W., 1992. *Die Transzendente Deduktion der Kategorien in der ersten Auflage der Kritik der reinen Vernunft*, Frankfurt: Klostermann.
- Cassam, Q., 1999. *Self and World*. Oxford: Oxford University Press.
- Chignell, A., 2010. "Causal Refutations of Idealism," *Philosophical Quarterly*, 60: 487–507.
- Dicker, G., 2004. *Kant's Theory of Knowledge*, New York: Oxford University Press.
- Dicker, G., 2008. "Kant's Refutation of Idealism," *Noûs*, 42: 80–108.
- Dickerson, A. B., 2004. *Kant on Representation and Objectivity*, Cambridge: Cambridge University Press.
- Dyck, C., 2011. "Kant's Transcendental Deduction and the Ghosts of Descartes and Hume," *British Journal for the History of Philosophy*, 19: 473–96.
- Förster, E. (ed.), 1989. *Kant's Transcendental Deductions*, Stanford: Stanford University Press.
- Grüne, S., 2011. "Is There a Gap in Kant's B Deduction?" *International Journal of Philosophical Studies* 19: 465–90.
- Guyer, P., 1977. "Review of W. H. Walsh, *Kant and the Criticism of Metaphysics*," *Philosophical*

*Review*, 86: 264–70.

- Guyer, P., 1987. *Kant and the Claims of Knowledge*, Cambridge: Cambridge University Press
- Hanna, R., 2008. “Kantian Non-Conceptualism,” *Philosophical Studies*, 137: 41–64.
- Hanna, R., 2011. “Beyond the Myth of the Myth: A Kantian Theory of Non-Conceptual Content,” *International Journal of Philosophical Studies*, 43: 91–110.
- Heidegger, M., 1929/1997. *Immanuel Kant und das Problem der Metaphysik*, Friedrich Cohen; translated as *Immanuel Kant and the Problem of Metaphysics*, by Richard Taft, fifth edition, Bloomington: Indiana University Press.
- Henrich, D., 1968–9. “The Proof-structure of Kant's Transcendental Deduction,” *Review of Metaphysics*, 22: 640–59.
- Henrich, D., 1989. “Kant's Notion of a Deduction,” in Förster 1989, pp. 29–46.
- Howell, R., 1992. *Kant's Transcendental Deduction*, Dordrecht: Kluwer.
- Hume, D., 1739. *A Treatise of Human Nature*, Oxford: Oxford University Press, 1978.
- Hume, D., 1748. *An Enquiry Concerning Human Understanding*, Oxford: Oxford University Press, 2005.
- Keller, P., 1998. *Kant and the Demands of Self-Consciousness*, Cambridge: Cambridge University Press.
- Kemp-Smith, N., 1923. A Commentary to Kant's *Critique of Pure Reason*, New York: Humanities Press, 1962, second edition.
- Kitcher, Patricia, 1990. *Kant's Transcendental Psychology*, New York: Oxford University Press.
- Kitcher, Patricia, 2011. *Kant's Thinker*, New York: Oxford University Press.
- Kitcher, Philip, 1981. ‘How Kant Almost Wrote ‘Two Dogmas of Empiricism,’ *Philosophical Topics*, 12: 217–49.
- Kuehn, M., 1997. “The Wolffian Background to Kant's Transcendental Deduction,” in *Logic and the Workings of the Mind*, P. A. Easton (ed.), Atascadero: Ridgeview Press, pp. 229–50.
- Longuenesse, B., 1998. *Kant and the Capacity to Judge*, Princeton: Princeton University Press.
- Longuenesse, B., 2000. “Kant's Categories and the Capacity to Judge,” *Inquiry*, 43: 91–110.
- McCann, E., 1985. “Skepticism and Kant's B-Deduction,” *History of Philosophy Quarterly*, 2: 71–89.
- McLear, C., 2014. “Two Kinds of Unity in the Critique of Pure Reason,” *Journal of the History of Philosophy*, forthcoming
- Melnick, A., 1973. *Kant's Analogies of Experience*, Chicago: University of Chicago Press.
- Nagel, T., 1997. *The Last Word*, Oxford: Oxford University Press.
- Paton, H. J., 1936. *Kant's Metaphysic of Experience*, London: Allen and Unwin, 2 vols.
- Pereboom, D., 1990. “Kant on Justification in Transcendental Philosophy,” *Synthese*, 85: 25–54.
- Pereboom, D., 1995. “Self-understanding in Kant's Transcendental Deduction,” *Synthese*, 103: 1–42.
- Pereboom, D., 2006. “Kant's Metaphysical and Transcendental Deductions,” in *The Blackwell Guide to Kant*, G. Bird (ed.), Oxford: Blackwell, pp. 154–68.
- Pollok, K., 2008. “‘An Almost Single Inference’ – Kant's Transcendental Deduction of the Categories Reconsidered,” *Archiv für Geschichte der Philosophie* 90: 323–45.
- Rorty, R., 1979. “Transcendental Arguments, Self-Reference, and Pragmatism,” in Bieri, Horstmann & Krüger (eds.) 1979, pp. 77–103.
- Russell, B., 1912. *The Problems of Philosophy*, London, Williams and Norgate; New York, Henry Holt and Company.
- Sacks, M., 1999. “Transcendental Arguments and the Inference to Reality,” in Stern 1999a, pp. 67–83.
- Sacks, M., 2000. *Objectivity and Insight*, Oxford: Oxford University Press.

- Schulking, D., 2012. "Kant, Non-Conceptual Content, and the 'Second Step' of the B-Deduction," *Kant Studies Online*: 51–92.
- Schulking, D., 2013. *Kant's Deduction and Apperception*, Basingstoke, UK: Palgrave Macmillan.
- Smit, H., 1999. "The Role of Reflection in the Critique of Pure Reason," *Pacific Philosophical Quarterly*, 80: 203–23.
- Smit, H., 2009. "Kant on Apriority and the Spontaneity of Cognition," in *Metaphysics and the Good: Themes from the Philosophy of Robert Merrihew Adams*, Oxford: Oxford University Press, pp. 188–251.
- Stern, R. (ed.), 1999a. *Transcendental Arguments*, Oxford: Oxford University Press.
- Stern, R., 1999b. "On Kant's Response to Hume: The Second Analogy as Transcendental Argument," in Stern 1999a, pp. 47–66.
- Stern, R., 2000. *Transcendental Arguments and Scepticism: Answering the Question of Justification*, Oxford: Oxford University Press.
- Strawson, P. F., 1966. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*, London: Methuen.
- Stroud, B., 1968. "Transcendental Arguments," *Journal of Philosophy*, 65: 241–56.
- Stroud, B., 1994. "Kantian Argument, Conceptual Capacities, and Invulnerability," in Paolo Parrini (ed.), *Kant and Contemporary Epistemology*, Dordrecht: Kluwer, pp. 231–51.
- Stroud, B., 1999. "The Goal of Transcendental Arguments," in Stern 1999a, pp. 155–72.
- van Cleve, J., 1999. *Problems From Kant*, Oxford: Oxford University Press.
- Wolff, R. P., 1963. *Kant's Theory of Mental Activity*, Cambridge: Harvard University Press.
- Vaihinger, H., 1881–2. *Commentar zu Kants Kritik der reinen Vernunft*, Stuttgart: W. Spellman.
- Vogel, J., 1993. "The Problem of Self-Knowledge in Kant's 'Refutation of Idealism,'" *Philosophy and Phenomenological Research*, 53: 875–87.

## Academic Tools

-  [How to cite this entry.](#)
-  [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).
-  [Look up this entry topic](#) at the [Indiana Philosophy Ontology Project](#) (InPhO).
-  [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

## Other Internet Resources

- [Kant on the Web](#), including many links to other sites and e-texts, maintained by Stephen Palmquist.

## Related Entries

[Kant, Immanuel](#) | [Kant, Immanuel: account of reason](#) | [Kant, Immanuel: aesthetics and teleology](#) | [Kant, Immanuel: and Hume on causality](#) | [Kant, Immanuel: and Hume on morality](#) | [Kant, Immanuel: and Leibniz](#) | [Kant, Immanuel: critique of metaphysics](#) | [Kant, Immanuel: moral philosophy](#) | [Kant, Immanuel: philosophical development](#) | [Kant, Immanuel: philosophy of religion](#) | [Kant, Immanuel:](#)

[philosophy of science](#) | [Kant, Immanuel: social and political philosophy](#) | [Kant, Immanuel: theory of judgment](#) | [Kant, Immanuel: view of mind and consciousness of self](#) | [Kant, Immanuel: views on space and time](#)

[Copyright © 2013](#) by  
Derk Pereboom <[dp346@cornell.edu](mailto:dp346@cornell.edu)>

[Open access to the SEP is made possible by a world-wide funding initiative.](#)  
[Please Read How You Can Help Keep the Encyclopedia Free](#)

**Stanford** | Center for the Study of  
Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2014](#) by [The Metaphysics Research Lab](#),  
Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054