What is the ground of the relation of that in us which we call "representation" to the object?...Our understanding, through its representations, is not the cause of the object (save in the case of moral ends), nor is the object the cause of the intellectual representations in the mind (in sensu reali). Therefore the pure concepts of the understanding must not be abstracted from sense perceptions, nor must they express the reception of representations through the senses; but though they must have their origin in the nature of the soul, they are neither caused by the object nor bring the object itself into being. 1

In the letter to Marcus Herz from which this passage is taken, Kant says that the question of the ground of the relation of representation to its object "constitutes the key to the whole secret of hitherto still obscure metaphysics." The obscure metaphysical question which occupied Kant was the problem of understanding how knowledge is possible. Knowledge is a matter of our representing the world correctly. But if mind and world are separate, how can we know that the mind represents the world correctly? In response to the conflict between rationalism, which dogmatically asserts that the mind has knowledge of the world, and empiricism, which leads to skepticism about our knowledge of the world, Kant insisted we must first understand how the soul grounds the relationship between our representations and their objects.

While philosophy has since become uneasy about the soul, the thesis that the nature of the mind is crucial to metaphysics has been an enduring Kantian legacy. For Kant's conception of the soul includes two major contributions to our understanding of the mind. First, he decisively rejected any conception of the mind as consisting of objects: representation is a rule-governed activity, and an account of representation must consist in an analysis of those rules. Second, the rules which constitute the mind also constitute the world. Mind and world are so interrelated that they must either be understood together or not at all.

But what is the relationship between such programmatic commitments and a specific theory of mental functioning? Kant's exposition of his Copernican revolution is embedded in an aprioristic psychological theory, a theory that has since been discredited. This forces us to make a choice: either we ensure that Kant's insights are put on a philosophical basis that is independent of psychological theory, or we integrate them into empirical psychology. While Kant scholars have explored the first alternative, most philosophers of mind have embraced the second. In this paper, I approach the vexed question of the relationship between empirical psychology and the view of the mind which we have inherited from Kant by way of a discussion of Kant's argumentative strategy in the Transcendental Deduction.

In the Critique of Pure Reason, Kant set himself the task of providing a philosophical justification of our claim to have knowledge of the world. His justification relies on a distinction between three types of argument about the relationship between mind and world: empirical, dogmatic, and transcendental arguments.

"Empirical" reasoning makes use of empirical statements about the origin of our knowledge. Kant called such an argument an "empirical deduction." It "shows the manner in which a concept is acquired through experience and through reflection upon experience, and...therefore concerns, not its legitimacy, but only its de facto mode of origination." 2 So while it may tell us how we arrived at our present beliefs, it cannot show that they are true.

"Dogmatic" reasoning analyzes the logical relations that hold between our beliefs. It can lead to conclusions which are logically necessary, but its conclusions are analytic. It can only tell us of the relations that hold between our thoughts, while what we need is to show that certain thoughts about the world are true. "For such a purpose, the analysis of concepts is useless, since it merely shows what is contained in these concepts, not how we arrive at them a priori." 3

Kant concluded that neither "empirical" nor "dogmatic" reasoning could bridge the gap between mind and world. Dogmatism is particularly instructive here, for it assumes we can make use of the principles we apply in experience.
without first investigating in what way and by what right reason has come into possession of these concepts. Dogmatism is thus the dogmatic procedure of reason, without previous criticism of its own powers.

The dogmatist will reply that it is entirely reasonable to start from the standards we ordinarily employ. But to agree with common sense is not to vindicate it: we need a justification of the standards we ordinarily employ. An “appeal to the common sense of mankind [is] an expedient which always is a sign that the cause of reason is in desperate straits,” and so Kant holds that the problem of knowledge calls for a critique of pure reason. We must understand the power and limits of pure reason before we can say what it can and cannot achieve. One can also speak of answering the skeptic: Kant calls the skeptic “the taskmaster who constrains the dogmatic reasoner to develop a sound critique of the understanding and reason.”

Kant’s third mode of argument, the transcendental or critical argument, is designed to reach substantive conclusions about matters which skepticism doubts while accepting its starting point. In such an argument “the game played by idealism has been turned against itself.” For a critical argument aims to show that skeptical doubt depends on accepting conditions that skepticism claims to repudiate. Jonathan Bennett calls this skeptical point of departure the “Cartesian Basis,” the “intellectual situation in which one attends to nothing but one’s mind and its states.” Kant, like Bennett, thinks this is the correct starting point in philosophy:

While... the skeptical procedure cannot of itself yield any satisfying answer to the questions of reason, none the less it prepares the way by arousing reason to circumspection, and by indicating the radical measures which are adequate to secure it in its legitimate possessions.

So, a transcendental argument starts from a characterization of immediate experience and shows that this characterization commits one to more than the skeptic thinks. This calls for

(a) A characterization of some aspect a of experience.

(b) An argument that a is not possible unless some condition b obtains.

The skeptic may, of course, challenge our choice of a, arguing that there could be experiencing subjects whose experience does not include a. For example, Strawson gives an argument which starts from the premise that a subject must be able to distinguish those states that are one’s own and those that are not. The skeptic can reply that we can imagine subjects that do not satisfy such high standards of self-awareness. Indeed, there are many possible values for a: it has even been argued that we can imagine experience in a spaceless and timeless world.

This suggests that we need to add to our argument schema

(c) An argument that b is also a condition for every characterization of experience.

But this move does not so much fix up the preceding argument as open it up to further objections. How, for instance, could we be sure that we had exhausted all the possible alternatives? To meet this difficulty, we must look for aspects of experience that are essential. If we had an analysis of the necessary conditions for something’s being an experiencing subject, then the problems (c) was designed to remedy could not arise. We would then have the following argument schema:

(A) An argument that possession of characteristic a is a necessary condition for something’s being an experiencing subject.

(B) An argument that b is a necessary condition for a.

If we can find a pair of arguments that fit this schema, we will have a transcendental argument. To meet the skeptic’s demands, a must be an aspect of experience which even the skeptic cannot dispense with; to fully answer the skeptic, b must affirm what the skeptic denies. For instance, it might be a claim such as “I am part of a spatiotemporally extended world” or “I am part of a community with a shared language.”

Let us look at Kant’s reasons for thinking that a certain conception of experience must be accepted by the skeptic, part (A) of the argument schema. Kant holds that philosophy deals with the limits of possible experience, where “experience” is taken to be the conceptualized experience of a judging subject. But can’t we be conscious without judging, without being self-conscious? Imagine being in an accident in which one suffered such severe injuries that one was kept alive by machinery, unable to speak or think. With the disappearance of the ability to think and judge, there would be no “I,” no judging subject. All that would be left would be a body on a life-support system. Something which is conscious and has sensations but has no conceptualized experience would be “merely a blind play of representations, less even than a dream”—there would be no judgment, and hence no subject. How is Kant to
establish this distinction between mere sensation and experience proper? First, it must be conceptualized experience, experience in which concepts are applied in making judgments. The skeptic will grant that judgments of the form "It seems to me that p" are made within the Cartesian Basis but does not see this as a concession to Kant. Kant's reply is that judgment has a larger role than the skeptic admits. While most experience does not involve explicit judgment, judgment enters into all experience, for all experience is potentially judgeable. As Kant cryptically puts it, "it must be possible for the 'I think' to accompany all my representations."15

Second, it must be the experience of a judging subject. Although I do not "perceive something simple and continu'd" which is my "self," my experiences are my experiences only because they stand in certain relations to each other. These connections are what constitute the synthetic unity which obtains between a person's experiences. For

As my representations (even if I am not conscious of them as such) they conform to the condition under which alone they can stand together in one universal self-consciousness, because otherwise they would not all without exception belong to me.17

The relations between our representations lead us to think of them as coordinated toward that focus imaginarius which is the transcendental ego.

In the Transcendental Deduction, Kant sets himself the task of showing that the formal aspect of experience which makes it a subject's experience can be identified with the formal aspect which makes it judgeable, that the two conditions I have outlined—judgeability and subjectivity—are ultimately one and the same. But how is this claim to be established? It seems to say that for a representation to be my representation, I must be able to judge that it is mine. Indeed, some commentators have alleged that this thesis rests on an equivocation. The charge can be set out as follows. Kant starts from the trivially analytic

(1) All my representations are my representations.

This leads to the equally uncontroversial claim that

(2) All my representations must be such that they satisfy the conditions in virtue of which they are my representations.

Kant then slips without justification to the much stronger claim:

(3) All my representations must be such that I can judge them to be my representations.

But, his critics respond, if I can have a long-lost cousin without being able to recognize her as my cousin, why can't I have a representation that I can't recognize as mine?18

One cannot defend Kant by saying that (3) is part of what he means by a representation, or that he defines a representation as judgeable. For this would be a retreat to dogmatism and the skeptic will reply that no reason has been given for it. If (3) is to be justified, Kant must concede that an unjudgeable representation is a logical possibility; and so he must consider the possibility that

there might exist a multitude of perceptions, and indeed an entire sensibility, in which much empirical consciousness would arise in my mind, but in a state of separation, and without belonging to a consciousness of myself.19

In that case the manifold of sense would lack "affinity," and the mind would be unable to unify it. What, then, would happen if the mind were to be presented with such an unsynthesizable manifold? Kant's answer is that as the mind cannot unify it, it would be "nothing to me."20 For a representation to be a thought of mine, it must stand in whatever relation to my other thoughts which is constitutive of my identity. If I cannot judge it, then it cannot enter that circle. The skeptic's conception of a subject that only judges part of the time rests on an inadequate conception of the self: for experience to be assignable to a self at all, it must be at least potentially judgeable. In the B Deduction Kant says that

The thought that the representations given in intuition one and all belong to me, is therefore equivalent to the thought that I unite them in one self-consciousness, or can at least so unite them; and although this thought is not itself the consciousness of the synthesis of the representations, it presupposes the possibility of that synthesis. In other words, only in so far as I can grasp the manifold of the representations in one consciousness, do I call them one and all mine.21

From this vantage point, the outline of Kant's path to the thesis that the categories must be applied in all experience is clear. All experience must be judgeable, it must be possible to attach an "I think" to it, and so the manifold must be such that it can be combined and so brought to the unity of apperception.

For without such combination nothing can be thought or known, since the given representations would not have in common the act of the apperception "I think," and so could not be apprehended together in one self-consciousness.22
To judge is to apply the categories: if all intuition is judgeable, then the categories are applicable to all intuition.

Judgmental experience must have a certain formal structure, and this structure, it turns out, is what makes it subjective. Whatever differences there may be between any two subjects, they must have this structure in common. This, then, is Kant's justification for the selection of a, that aspect of experience on which his argument is to turn. Regardless of the gradations between the most sophisticated consciousness and the most rudimentary, we can draw a sharp line between possible conceptions of experience by means of the criterion of self-conscious experience: is the experience the conceptualized experience of a judging subject? This criterion does not force us to say that creatures that do not make judgments cannot have experience. It only requires that we sharply distinguish two kinds of experience.

Although we can raise the bare possibility of an alternative synthesis of the manifold, philosophical reflection has led us to see that it is not a real possibility, for it is ruled out by the train of thought we have just examined. This thinning out of the skeptic's speculative psychologies has its limits, however. For instance, it still leaves open the possibility that there might be subjects which satisfied Kant's criterion but whose experience was phenomenologically quite different. For the criterion is metaphysical: it tells us what conditions something must satisfy for it to be a subject. It does not, in itself, give any answer to related epistemological questions about identifying subjects, or what things are like for a subject. In this connection we might consider familiar examples, such as people who lack a sense or have heightened sensitivities, or more exotic alternatives, such as having eyes on two separate bodies, or the sensations of bats, which orient themselves by listening to the way in which the high-pitched sounds they emit are reflected back by the objects around them. The last cases are so different from our own that we may well be unable to imagine what such experience is like.

Kant does not trouble himself with such questions, perhaps because he thinks that the rejection of skepticism yields much more than the bare criterion of self-conscious experience. In his hands, the argument also leads to a detailed account of the operation of our mental faculties in synthesizing experiences. Once that account is in place, such vertiginous alternatives are no longer possible.

The train of thought I have just sketched suffices, I believe, to set out much of what is most powerful in Kant's account of the relationship between judgment and self-consciousness in the Transcendental Deduction. What role does the synthetic activity of the mind play in his argument? Do we really need an account of how it operates? For if we ignore his detailed theory, we can allow our reconstructed argument to rely on the bare thought that there are processes that underlie our experience and leave the task of specifying those processes to neurophysiology. Kant's insights into the nature of the mind might in this way be separated from his mistaken theory. But once we attempt to reconstruct his argument without his transcendental psychology, we have to ask ourselves whether the reconstructed argument can stand alone.

Kant not only tells a detailed psychological story about how the mind constitutes the world, he also makes continual use of it in the course of his argument. Indeed, in the first edition of the Critique of Pure Reason, he gave an elaborate description of the way in which the mind synthesizes the manifold of sense into sensory experience. Much of this transcendental psychology was eliminated in the second edition. In particular, the new exposition of the Transcendental Deduction leaves out most of the detailed description of the operation of the faculties to be found in the original version. The first-edition exposition gives the central role to the constitutive activity of the imagination; in the second, the concept of judgment occupies center stage.

Six years of further work on his system had given Kant a clearer idea of his argument. At first sight, it looks as though he had taken the opportunity to remove unnecessary commitments to faculty psychology. In support of this reading, one can point to §§15-20 of the B Deduction, which contains a very clear exposition of the line of argument I have just set out. The sections that follow make it equally clear that Kant thought of this as only a part of his overall argument. In §21 he tells the reader that in §20 "a beginning is made of a deduction of the pure concepts of understanding." So far, he says, he has abstracted from the spatiotemporal nature of human sensibility, an omission that must be made good before the proof can be completed. Only in §26 will the "a priori validity of the categories in respect of all objects of our senses" be made good.

The first part of the argument of the B Deduction concludes with §20, which is a brief and well-organized summary of the overall course of the argument. The title of the section is a statement of what Kant has to prove: "All sensible intuitions are subject to the categories, as conditions under which alone their manifold can come together in one consciousness." To see why he thinks the argument can only be completed in §26, let us look one by one at each sentence in §20.

Kant starts by reminding us that

The manifold given in a sensible intuition is necessarily subject to the original synthetic unity of apperception, because in no other way is the unity of intuition possible.

This amounts to a summary reminder of an argument we have just met: unless the manifold were synthesizable, it would be "as nothing" to me.
and hence could not be my experience. Next, he makes the point that synthesis is a matter of judgment:

that act of understanding by which the manifold of given representations (be they intuitions or concepts) is brought under one apperception, is the logical function of judgment.

Having brought these two main strands of his argument together, one might expect that Kant would bring in the relationship between judgment and the categories—

Now the categories are just these functions of judgment, in so far as they are employed in determination of the manifold of a given intuition.

—and so reach the desired conclusion:

Consequently, the manifold in a given intuition is necessarily subject to the categories.

Indeed, in §§15–19 of the B Deduction, this is what he has already done. However, between the first two sentences of §20 and the last two, Kant inserts his caveat:

All the manifold, therefore, so far as it is given in a unified empirical intuition, is determined in respect of one of the logical functions of judgment, and is thereby brought into one consciousness.

This raises anew the possibility of an unsynthesizable manifold being given to sensibility. For if there were a manifold of sense that could not be "given in a unified empirical intuition," it need not be "determined in respect of any of the logical functions of judgment." The preceding argument has shown that whatever is given in a unified empirical intuition—immediate experience—must be judgeable. This is insufficient, for it leaves open the possibility that the sensory material which is initially presented to the mind may be so chaotic that it cannot be synthesized into a unified empirical intuition.

Kant's caveat makes it quite clear that he would not have been content with the logical reconstruction of his deduction that I gave earlier. For in abstracting from any consideration of the precise nature of synthesis, it can have nothing new to say in reply to the suggestion that our initial sensory input, rather than the content of sensory experience itself, might be unjudgeable. It can only reply that both suggestions are equally inconceivable, for in either case we are talking about experience which would be unjudgeable and so "as nothing" to me.

In §26, Kant says that the first half of the argument only shows that the categories are applicable in all knowledge of "objects of an intuition in general." This leaves the task of showing that "whatever objects may present themselves to our senses" can be brought under the categories. In other words, he still has to show that every manifold of sense must be unified by a synthesis which employs the categories and cannot merely be collocated by empirical laws of association.

Now we can see the force of the remark in §21 that the preceding argument involves abstracting from the spatiotemporal nature of our sensibility. The argument leading up to §20 draws out the connections between the subject and object of experience in a wholly general way, while the subsequent argument will turn on considering the fact that experience is spatiotemporal. Kant's account of the spatiotemporal constitution of experience provides the basis for the second half of his argument.

Kant's solution turns on an appeal to the synthetic activity of the mind. Space and time are the forms of all our sensibility, so everything that is given to us in sensibility is presented under at least one of these forms. The Transcendental Aesthetic has shown that we also have a priori intuitions of space and time. Any intuition presupposes the existence of a synthesis which has brought about the unity of that intuition. As the categories are applied in constituting the intuitions of space and time, which in turn contain the entire manifold of sense, all intuition presupposes a synthesis by means of the operation of the understanding. But if the categories are applied to the entire manifold of sense in constituting space and time, then the categories must be applicable to everything given in that manifold.

Even if Kant never deviated from considering transcendental psychology an indispensable part of his argument, this hardly settles the question whether he was right to do so. The problem he tries to solve in the second half of the B Deduction only arises once one accepts that there is a need for an account of the constitution of sensory experience. The solution to his problem is ad hoc, too. It requires a substantial and unjustified revision of the introductory account of space and time given in the Transcendental Aesthetic—Kant summarily asserts, contrary to the Aesthetic, that formal intuition presupposes a synthesis.

If the problem and the solution are equally spurious, why should we let this confusion worry us? At the very least, it is a sign that Kant felt an acute need for an account of the processes in the mind which make judgment possible. Further, it points to an underlying difficulty: while Kant spurned opinions and hypotheses, he nevertheless made claims about the mind which we now judge to be false, insofar as we can understand them at all. Lacking his unquestioning acceptance of faculty psychology, we will say that it is not a description of the nature of the
mind and that its true locus is the Kantian texts themselves. If we put aside the exegetical exercise of trying to understand the role of transcendental psychology in those texts, we are left with the question of what we can salvage from the Kantian program.

Kant says that transcendental arguments can only tell us about the constitution of the phenomenal world, not things in themselves. Strawson and most subsequent Kant commentators advise us to disregard transcendental idealism as a part of the metaphysical strand of the Critique which has not stood the test of time. Talk of the mind constituting the phenomena is to be rejected as part of Kant's psychologistic tendencies. On this "austere" reconstruction, transcendental arguments simply establish how things must be if experience is possible.

In this spirit, we might try to reconstruct a transcendental argument along the lines I sketched out earlier. Such an argument moves from \( a \), a characterization of the experiencing subject, to \( b \), the conditions that must obtain if there is to be such experience. If the argument is valid, then the only way to avoid its conclusion is to reject the premise. To do that is to hold that \( a \) is not essential to experience. Kant and Strawson charge that the skeptic has been caught out. To reply, the skeptic would have to show that we can dispense with the subject. But the skeptic's starting point, the Cartesian Basis, is the situation in which one attends to nothing but one's mind and its states.

Richard Rorty had described such an argument as an "anti-parasite" argument. It moves from the chosen \( a \) to some condition of its employment. Should the skeptic set out a conception of experience that lacks \( b \) but still contains \( a \), then one can show that the conception is not self-sufficient, being "parasitic," for skepticism must also accept \( a \). Such an argument is only a standing challenge to the skeptic to find an autonomous conception of experience, one that lacks \( a \) and so lacks \( b \). In any case, the skeptic is unlikely to accept the claim that transcendental arguments establish how things must be. Instead, he or she will argue that even a successful transcendental argument would only show how we must think of the world, not how it is in itself, and so leaves open the possibility that we must think of it wrongly.

One might think that the verification principle could be used to bridge the gap: if a statement's meaning consisted in a method for ascertaining its truth, to understand \( b \) is to be able to tell if it is true. However, the verification principle is of little use against the skeptic, who can argue that \( b \) is meaningless rather than false. In any case, a verification principle strong enough to underwrite such an argument could only make the argument valid at the price of making it obviously unsound.

If a transcendental argument is not the final answer to the skeptic, but only one move in a continuing struggle, then it no longer has the decisive power Kant promised. In that case, we may well return to the form of argument I mentioned briefly toward the beginning of this paper: empirical reasoning. The empirically minded philosopher thinks that some principles must underlie the functioning of the mind, and that psychology, neurology, or artificial intelligence will eventually enable us to understand them. Many contemporary philosophers of psychology have turned to such theories. They believe that science will set out the rules on which the mind operates and so enable us to understand the mind and its place in nature. However, most formulations of this approach depend on assumptions about the nature of the mind that are as unjustified as any to be found in dogmatism or transcendentalism. For they assume that there are processes in the brain which can be correlated with mental processes. Yet we have no neurophysiological theory that actually does this. Why should we expect that theory will in due course arrive at some such correlation between terms referring to mental tokens and terms referring to physical tokens? Current work on connectionist theories of neural functioning strongly suggests that no such principles will be found.

Perhaps the questionable move in the Kantian argument about the ground of representation on page 220 (this volume) is the first inference, the one we took for granted. Starting from the premise that "All my representations are my representations," we inferred that all my representations must satisfy whatever the conditions are for them to be my representations. In making that move, both Kant and the empiricist tacitly assume that the psychological phenomena are grounded in an underlying order. But "why should there not be a psychological regularity to which no physiological regularity corresponds? If this upsets our concept of causality then it is high time it was upset." We have simply taken for granted the assumption that the relationship between our representations and their objects must be grounded in a certain kind of psychological theory. Like the Ptolemaic belief that the sun goes around the earth, Kant's faculty psychology, or the phrenologists' account of the nature of the mind in terms of the size and shape of the skull, this is a natural interpretation, an inexplicit assumption, not a scientific result. Once we see that it is an assumption, we may wonder why we were so convinced it was true.

NOTES

Part Four: Kant


5. A783–84 = B811–12.
6. A769 = B797.
10. This anti-Kantian strategy is the theme of chap. 4 of John Mackie's The Cement of the Universe (Oxford: Oxford University Press, 1974).

11. Peter Strawson, Individuals (Methuen, 1959), chap. 2.
17. B132–33.
18. On pp. 83–84 of On What There Must Be (Oxford: Oxford University Press, 1974), Ross Harrison puts this objection very clearly, but he seems to consider it obvious enough that he does not bother to argue for it.


19. A122; cf. A89/B122ff. The passage continues with the words: "This, however, is impossible." We have yet to see how this can be shown.
22. B137.

Wittgenstein thought that visual experience with two bodies was inconceivable and that this pointed to an important epistemological asymmetry, an asymmetry between self-knowledge and knowledge of others:

What would it be like if I had two bodies, i.e. my body were composed of two separate organisms?

Here again, I think, we see the way in which the self is not on a par with others, for if everyone else had two bodies, I wouldn’t be able to tell that this was so.

Can I imagine experience with two bodies? Certainly not visual experience.


24. B144.
26. B143. I have omitted Kemp Smith's capitalization of the nouns. All of §20 is on p. 143 of the B edition.
Imagine the following phenomenon. If I want someone to take note of a text that I recite to him, so that he can repeat it to me later, I have to give him paper and pencil, and while I speak, he makes lines, marks, on the paper; if he has to reproduce the text later he follows those marks with his eyes and recites the text. But I assume that what he has jotted down is not writing. It is not connected by rules with the words of the text; yet without these jottings he is unable to reproduce the text; and if anything in it is altered, if part of it is destroyed, he gets stuck in his "reading" or recites the text uncertainly or carelessly, or cannot find the words at all.—This can be imagined!—What I called jottings would not be a rendering of the text, not a translation, so to speak, in another symbolism. The text would not be stored up in the jottings. And why should it be stored up in our nervous system?

The conclusion is more striking in the German: "stored up" translates "niederlegen," which can also mean to lay down, or to write down. I discuss this passage at much greater length in a paper on "Models of Memory: Wittgenstein and Cognitive Science."

36. I would like to thank Simon Blackburn, Patrick Maher, Bob Brandom, Janet Broughton, Hans Sluga, Elizabeth Calihan, Bert Dreyfus, and Nancy Mullenax for their comments on previous drafts of this paper.