

# The Role of Self-Knowledge in the *Critique of Pure Reason*

RICHARD F. H. POLT  
University of Chicago

In the *Critique of Pure Reason* Kant attempts, among other things, to solve two problems about our knowledge of the world. First, how can we know any necessary truths about the world, such as the principle that every event must have a cause? Kant follows Hume in holding that unless we are able to know such necessary truths, and to justify them without recourse to experience, we will not even be able to know any contingent truths about the world, for these all make use of the necessary principles. Second, how can I know that things other than I exist at all? This radical doubt about my knowledge of the world had, of course, plagued philosophers ever since Descartes. Kant's strategy for dealing with both these problems is to repudiate the assumptions behind earlier modern epistemology, and in particular, the kind of distinction that Descartes and Hume had made between self-knowledge and our knowledge of 'outer' things. For Kant, this distinction as it stands is too simplistic, and it inevitably leads to skepticism about our knowledge of the world. But he still distinguishes self-knowledge from other kinds of knowledge; in fact, self-knowledge plays a crucial role in Kant's new epistemology. Kant's innovation is to distinguish sharply between two sides of what his predecessors understood as self-knowledge: 'inner sense' and 'apperception.' This paper will show how Kant's conception of 'apperception' is meant to overcome Humean skepticism, and how 'inner sense' is meant to overcome Cartesian skepticism; I will then touch upon a complication in Kant's view of the relation between apperception and inner sense, a complication which relates directly to the problem of the 'thing in itself.' I limit myself to the Aesthetic and Analytic of the first *Critique*; my aim is not to criticize Kant's views, nor to compare his sometimes equivocal formulations, but to give a coherent summary to which I believe Kant would subscribe.

First, we must understand how Descartes' and Hume's problems arise, and how they relate to the question of the nature of self-knowledge. Descartes' *Meditations* typify the view of self-knowledge that characterized most pre-Kantian epistemology. For Descartes, I have immediate experience only of myself and what is 'in' me; from the sense-impressions, or appearances, in my subjective sphere, I then infer that 'outer' objects, or things in themselves, exist. This position clearly invites me to doubt whether such inferences about things outside me can ever be rationally justified; Descartes' own solution to this doubt relies on the goodness of God. For Descartes, self-knowledge, or the knowledge that I exist, is the only knowledge that is certain to begin with.

Hume's skepticism takes a different tack: in *An Inquiry Concerning Human Understanding*, sections IV and V, he claims that when we judge that events are necessarily connected as causes and effects, these judgments are merely mental habits without any rational justification. He argues, on the one hand, that the principle of cause and effect cannot be justified independently of experience, for truths that hold independently of experience must be tautologies; and since it is possible to imagine a world that would *not* conform to the principle of cause and effect, this principle is no tautology. On the other hand, the principle cannot be justified by experience, because this reasoning would be circular: all judgments from experience already rely on the principle. Hume concludes that any claims we make about things in the world are unjustified. To put it otherwise: for Hume, we can have self-knowledge, for we can make justified claims about what *impressions* we are having and about how our mind tends to connect them; but any claim about real connections among things themselves is untenable. Once again, we have a distinction between self-knowledge and knowledge of things 'outside' us.

How, then, does Kant use his own conception of self-knowledge to refute Descartes and Hume? Before we can answer this question, we must see how Kant conceives of the self—and this in turn depends on how he conceives of knowledge. Kant's predecessors generally held that knowing was either an essentially active or an essentially passive process; but Kant draws a fundamental distinction within the very process of knowing. For the rationalists, true knowing was active, in that it deduced all its conclusions from indubitable principles; for the empiricists, true knowing was passive, in that it depended wholly on what was given to us through the senses. But for Kant, the process of knowing is both active and passive. Knowing has two poles, the knower and the known; the knower actively produces concepts, which it applies to the intuitions, or appearances, it passively receives—and thus it makes them known.

The dual nature of knowing implies that we can make valid claims to knowledge only if we do not exceed the bounds of experience, i.e. of "empirical knowledge" (B 147)<sup>1</sup>, which always includes both concepts and intuitions. As we have noted, in receiving intuitions the mind is passive, whereas in producing concepts the mind is spontaneous (A 51). Kant entitles the ability to receive intuitions the faculty of *sensibility*, and the ability to produce concepts the faculty of *understanding*.

Kant's conception of knowledge is what leads him to distinguish two senses of the 'self.' We have seen that Kant conceives of the self as the knower—the active pole of the process of knowing. But the self can also

---

<sup>1</sup>References are to pages in the first (A) or second (B) edition of the *Critique of Pure Reason*, depending on where the passage first appeared; the translation is that of Norman Kemp Smith.

be an object of knowledge—and this is to say that it can be an object of experience, an object of empirical knowledge.<sup>2</sup>

How does the self become an object of empirical knowledge? Kant distinguishes two kinds of sensibility: 'outer sense' and 'inner sense.' Through outer sense we perceive "objects as outside us" in space, and through inner sense "the mind intuits itself or its inner state" in time (A 22). But according to Kant's conception of experience, inner and outer sense are fundamentally similar: both my experience of myself and my experience of other objects must depend on my receiving intuitions to which I apply concepts. Like knowledge of outer objects, any knowledge of my mind must have sensible intuitions as its content; and self-knowledge becomes objective and well-founded in applying concepts to these intuitions. Thus, for Kant, knowledge of the self is neither more nor less certain and immediate than any other empirical knowledge. We have no 'intellectual intuition' of ourselves: intellectual intuition, or intuitive understanding (cf. B 135), would be a purely spontaneous knowledge, which would not depend on the ability to receive intuitions (B 68). To deny intellectual intuition of the self is, in traditional terms, to imply that the subjective and the objective spheres of knowledge are alike in kind: both depend on sensibility, and they establish truths not by penetrating to a thing in itself at all, but by connecting appearances under determinate rules.

What, then, of the self that is the active pole of the process of knowing? Is it, too, an object of experience? For Kant, this cannot be. He conceives of the knowing self not as an object of knowledge at all, but as an ability: I must always be able to preface any piece of knowledge by 'I think'—otherwise, the knowledge would not belong to *my* experience. This capacity for self-consciousness, or 'apperception,' is completely different from inner sense, for it yields no intuition and no experience. 'I think,' or 'I am I,' cannot be considered knowledge. What Descartes called the *cogito* turns out not to be an item of self-knowledge at all.

But apperception plays a role in the *Critique of Pure Reason* that is as important as the role of the *cogito* in the *Meditations*: it is on apperception that Kant founds his refutation of Humean skepticism. To understand this refutation, we must consider Kant's notion that sensibility and understanding have 'forms.'

---

<sup>2</sup>The ideas of reason include a further meaning of 'self' or 'soul'; the idea of the soul as simple substance can be used as a regulative principle in psychology (A 682). Such ideas yield no knowledge proper, but we should not forget the crucial importance of non-cognitive ideas for Kant's thought as a whole. After all, Kant circumscribes the domain of knowledge in order to safeguard the ideas of a moral self and of freedom.

For Kant, sensibility and understanding operate according to particular rules, or 'forms'; and these forms are what give us 'synthetic *a priori* knowledge'—truths that are necessary, but not tautological. In the human mind, sensibility takes the forms of space and time, and understanding takes the forms of twelve 'categories' (which Kant divides into four groups). Space, time, and the categories are '*a priori* concepts' (B xvii); they are the ways in which we constitute unified objects of experience that are both intuited and thought. *A priori* concepts, says Kant, are *conditions of possibility of experience*; that is, any object we can know empirically must conform to the rules of sensibility and understanding. If we can know *a priori* that these concepts must apply to our experience, then we can have certain and substantive knowledge, synthetic *a priori* knowledge, about experience in general.

If Kant is right, then Hume is mistaken in assuming that any necessary truths must be tautological; the principle of cause and effect is one of the categories, or synthetic *a priori* truths about experience. But how do we know that our understanding must have special rules? Why is it legitimate to expect all experience to conform to these rules, or categories? What function do these categories serve? Here is where Kant uses apperception to justify his epistemology.

Kant argues in his "Transcendental Deduction of the Categories" that it is the 'I think' that makes the categories apply necessarily and validly to all intuited objects. The 'I think' must be able to accompany all our representations; but it can do so only if these are brought together into a unitary experience—'my' experience. In fact, apperception is simply the unification of my experience; it has no content (B 135), but is, rather, a 'function,' that is, the unity of an act of synthesis (cf. A 68). The categories, then, are nothing but the ways in which I thus synthesize intuited objects, fixing them within a spatio-temporal whole. Since representations cannot belong to my experience unless they are unified, all my knowledge must conform to the rules by which I unify. It is in this sense that Kant can call apperception the "supreme principle of all employment of the understanding" (B 136), and can even call the understanding the "faculty of apperception" (B 134, note). The 'I think' does not yield any knowledge in itself; but it makes all our knowledge possible.

How does Kant's approach obviate Humean skepticism? Hume had argued that the principle of cause and effect could never be proved from experience, for it already underlies all arguments from experience. But neither can the principle be proved independently of experience, for it is not a mere tautological 'relation of ideas.' Kant uses apperception as the basis for showing that not all *a priori* truths are tautologies. Apperception, and thus thinking in general, would be impossible if we were not able to connect appearances in relations such as cause and effect. These relations are thus necessary—not as mere 'relations of ideas,' but as the glue that binds experience together and makes it *someone's* experience.

The way Kant deals with the second problem about our knowledge of the world is still more complex. Kant's answer to Descartes' skepticism about the existence of objects 'outside' me in general depends not only on apperception, but on the relation between apperception and inner sense. We have seen that the two 'selves' in Kant's epistemology must be kept distinct: the empirical self differs from the transcendental 'I' of apperception. We fall into an illusion, a 'paralogism' of pure reason (B 406ff.), if we treat apperception as an object of knowledge, and equate the function 'I think' with a soul about which we can make judgments. But though the empirical and transcendental selves are distinct, they nevertheless stand in a complex relation. We must now consider the details of how apperception interacts with inner sense when we apply the categories to intuitions; we can then understand Kant's doctrine that inner sense depends on outer sense, which constitutes his refutation of Cartesian skepticism.

If we are to apply the categories to empirical objects, our mind must interpret the categories in terms of the particular forms of sensibility through which we intuit empirical objects. In other words, we must mediate between understanding and sensibility; there are rules for this mediation, rules which Kant calls 'schemata.' In his chapter on the schemata, Kant emphasizes the importance of time, the form of inner sense. For Kant, time is, as it were, the interpretive tool that inner sense provides us so that we can carry out the task set by apperception—the task of unifying our experience. The schemata, then, are procedures for forming 'images' of objects (A 140) which interpret the categories in terms of the objects' relation to time or their interrelation within time. Thus, in applying categories of 'quantity' to an object I count, successively adding one homogeneous unit to another and thus "generating time itself" (A 143). Categories of 'quality' are schematized in terms of "filling" time with a degree of sensation (A 145). Categories of 'modality,' including existence, are schematized in terms of "whether and how an object belongs to time" in general (whether it can enter our experience, and when) (A 145). Lastly, categories of 'relation' are schematized in images of objects as existing at different determinate times, at the same determinate time, or throughout all time. This last schema, which interprets the category of substance in terms of *permanence*, is crucial, for it provides a stable object in reference to which all other objects can be temporally determined (A 143 f.).<sup>3</sup> But

---

<sup>3</sup>I shall not dwell on these details of the schematism, but we can note that Kant's view of time is rather simple: time is a sequence of homogeneous moments which I generate through counting, and objects are either there or not there at each moment. A critic of Kant could question whether this account of time is sufficient to characterize our experience of the self and the world.

here, inner sense needs outer sense: in order to complete the task of unification, we require not only time but space—and this will form the basis of Kant's refutation of Descartes.

As our next step towards understanding Kant's argument against Cartesian skepticism, we must look more closely at what Kant means by 'determination,' and at his claim that time alone is insufficient for this determination. Kant refers to the schemata as "*a priori* determinations of time" (A 145). The unity of apperception requires these schemata in order to effect "the determination of appearances in space and time in general" (B 168f.). We use the schematized categories to *determine* objects within a coherent experiential whole; in other words, in order to unify our experience, we must assign to each element of experience a definite moment within a single temporal sequence, and a definite place within a single spatial structure.

Now, we might ask whether intuitions do not already present themselves as unified and determined, thanks merely to the fact that inner sense has a particular form, namely time. Is *spatial* determination also necessary for coherent experience? Is it not enough to fit our experiences into a temporal sequence, a 'stream of consciousness'? Anticipating this thought, Kant stresses that *inner sense alone is incapable of determining objects*: "Inner sense...contains the mere form of intuition, but without combination of the manifold in it, and therefore so far contains no determinate intuition" (B 154). Inner sense is simply a "limiting condition" on how I must intuit myself (B 159); it is not a sufficient condition for determining objects. Time indeed contains certain 'relations' of representations (A 33, B 67), but these relations are not ordered enough to let us have knowledge. The contents of time are "in constant flux" (B 291); "perceptions come together only in accidental order, so that no necessity determining their connection is or can be revealed in the perceptions themselves" (B 219). If Kant is right, all this means that if we attempted to consider our experience only as a temporal succession of perceptions in our mind, we would be isolating it from what makes experience objective, i.e. knowable and coherent.

But *why* does time alone fail to make our experience knowable? The answer is that for Kant, objects can be determined only in relation to something permanent, and permanence can be found only in space. Kant's first account of inner sense may suggest that some experiences—abstract meditations in my mind, perhaps—can take place without the medium of outer sense. But Kant also says that "all thought must, directly or indirectly, by way of certain characters [or criteria, *Merkmale*], relate ultimately to intuitions" (B 33). Inner intuitions alone do not present us with the 'criteria' that would allow us to relate our thoughts adequately to

---

them; inner intuitions cannot be determined under the categories, because determination requires sensible permanence (B 291). We might have thought that this sensible permanence was already given us by the schema of substance, as an object existing throughout all time; but Kant denies that time alone can present us with permanence. To be sure, says Kant, time itself is "non-transitory and abiding" as the form which inner sense must always take (A 143); "time, however, cannot itself be perceived" (A 177). An empty time cannot be sensed; it is unimaginable (though not inconceivable), because not time itself, but only things in time can enter our experience (A 188). Through inner sense, then, we sense only perceptions within time—and these are "in constant flux": "in inner sense no permanent intuition is to be met with" (B 292). The result is that we must relate inner sense to outer sense, in which there *can* be perceivable permanence; only then can we determine objects under the categories, for this involves fixing them in relation to such permanence. Objects in space are independent of the mere sequence of our mental states; they can thus provide us with lasting points of reference, landmarks to which we can apply 'criteria' in accordance with the schematized categories.

Kant cannot mean that we depend on perceiving some particular spatial object which is indestructible; his point is, rather, that it is only as belonging to an 'outer,' spatial world that objects can be treated as independent of our own momentary perceptions. And we *must* treat objects as persisting independently of our stream of consciousness, in order to use them as landmarks by which we measure a unified sequence of time. As we assign our experiences a place in a single sequence, we apply criteria to these landmarks, and use them as stable elements against which we can plot the sequence of our own, unstable perceptions. Objects in space allow us to measure and determine objects in general, and thus to make them compatible with the unity of apperception. Outer sense is hence a condition of possibility of all experience—even 'inner' experience.

We are now ready to understand Kant's "Refutation of Idealism"—by which he means "the problematic idealism of Descartes, which holds that there is only one empirical assertion that is indubitably certain, namely, that 'I am'" (B 274). Here Kant establishes that the knowledge of my own existence already implies "the existence of objects in space outside me" (B 275). For Kant, *existence is a category*; when we think that a thing exists, we schematize this category in terms of inner sense, and we think that the thing is present "in some determinate time" (A 145). But as we have seen, we cannot determine time unless we refer to something permanent, which can be intuited only in space. Hence "inner experience in general is possible only through outer experience in general" (B 278).

Here Kant has used his views of knowledge, the self, and self-knowledge to give a radically new sense to the 'inner and outer spheres' of his philosophical predecessors. For Kant, it is hopeless to attempt to know myself (or my 'inner' perceptions) independently of knowing other objects,

the objects perceived in these perceptions: unless I know objects in space, I cannot determine myself as a unified object, and thus I cannot even assert my own existence. In other words, it is an allusion to think that I can know about my 'representations' without knowing about any of the objects that I represent. I am immediately conscious of "a thing outside me and not...the mere representation of a thing outside me" (B 275). 'Inner sense' is indeed the sense by which I intuit myself and my inner state (A 22); but this intuition cannot rise to the status of knowledge unless it meshes with my 'outer sense' of things other than I. Kant still uses the words 'inner' and 'outer' here, spatial terms, simply because space is the form in which we are given objects other than ourselves; in Kant's hands, these terms no longer imply that we are locked into our consciousness as if into a prison cell.

We have now seen how Kant uses the two sides of self-knowledge--apperception and inner sense--to counter Humean and Cartesian skepticism. But in order completely to understand the role that self-knowledge plays in Kant, we must touch upon a complication--a point in the *Critique* that relates to deeper questions, questions that I believe Kant did not work out thoroughly.

Kant wishes to allow for an 'existence' prior to the category of existence, so that it can be said in some sense that the 'I' of apperception exists. He also speaks as if the distinction between transcendental and empirical self were founded in a deeper identity. Despite the difference between the mere thought "that I am" (B 157), which is unitary and purely spontaneous, and the knowable self which is determined in a complex world, the two selves are identified as "the same subject" (B 155): inner intuition is, after all, "self-intuition" (B 157, note). Because of this identification of the transcendental with the empirical self, we cannot say simply that the latter exists and the former does not; there must be a sense in which existence is "already given" (though not yet determined under the categories) in the 'I think.' "The 'I think' expresses the act of determining my existence. *Existence is already given thereby*, but the mode in which I am to determine this existence, that is, the manifold belonging to it, is not thereby given. In order that it be given, self-intuition is required" (B 157, note; my emphasis). Kant conceives of this existence given in the 'I think' as "an indeterminate empirical intuition"; it is "something real [i.e., something that is, and which "corresponds to a sensation in general," A 143] that is given...to thought in general"--but this real something is not given "as appearance," and "precedes the experience which is required to determine the object of perception through the category in respect of time" (B 422 f., note). The *Critique* insufficiently explores this vague conception of a "being itself [*das Wesen selbst*]" that is prior to any determinate mode of intuition (B 429). Yet the conception springs from the very idea of inner sense as *the mode in which we intuit ourselves*, for this implies an identity of the knower with the existent known, and thus an



existence of the knower--which nevertheless cannot yet be the determinate existence of the known.

This point brings us to the last issue I would like to consider: the relation between self-knowledge and the famous problem of the 'thing in itself.' We have seen that Kant's double conception of self-knowledge, as apperception and inner sense, springs from his double conception of knowledge itself, as knower and known. We can characterize both knower and known negatively: the knower as such can never be an object of knowledge, and the known is always an appearance, never a thing in itself. There is a fascinating symmetry here. We have just heard Kant insist that the knower is "the being itself"--an existence of which we are aware, yet about which we cannot make judgments, for Kant's doctrine of inner sense implies that the knower can know even itself "only as appearance, not as it is in itself" (B 156). He insists, symmetrically, that although we can know nothing about the thing in itself, we must at least think it, as that which appears in appearances (B xxvi f.). This leaves us with the question of how we are to think the relation of knower in itself to thing in itself. The unprovable suggestion of the first *Critique*, which is further explored in Kant's moral and aesthetic writings, is that the subject in itself and the object in itself can be the same.