

THE REFUTATION OF THE REFUTATION

William B. Innis
Marquette University

In the midst of his explication of the “Postulates of Empirical Thought,” Kant asserts that he has reached the appropriate location in the *Critique of Pure Reason* to launch into his somewhat polemical “Refutation of Idealism.” He begins this effort by drawing the distinction between two types of idealism which he views as troublesome (and these are obviously to be understood as something radically distinct from his own brand of idealism—namely, of the transcendental sort). These two particular types of idealism (which Kant identifies as “*material* idealism”) are *Problematic* idealism, which he attributes to René Descartes, and *Dogmatic* idealism, to which he attaches George Berkeley’s name. Refuting these types of idealism is something which is quite obviously important to Kant since throughout the body of his work eight such refutations appear (Walker, 71). This paper will focus on the particular aforementioned refutation which is exclusive to the second edition of the first Critique (and references to any of these other refutations will only be made insofar as they shed light on this argument in particular.)

The goal of this paper is threefold. First I shall provide an explication of the positions of *problematic* and *dogmatic* idealism, which will involve clarifying both their similarities and differences. Second I shall examine the argument in the “Refutation of Idealism” (as it is presented in the second edition of the first Critique) and other specific areas of the Kantian corpus which offer insight into this specific argument. The main thrust of this “Refutation of Idealism” turns on the necessary relationship between the empirical consciousness of one’s self and the consciousness of objects. Finally I offer a critical evaluation of this argument in order to underline the specific nature of the three deficiencies which it contains. In general I want to conclude that there are two counts of an unclear use of language employed by Kant (intentionally or unintentionally). This lack of clarity leads Kant both to assert two contradictory propositions about these “objects” of which one is conscious and about permanence in perception, and to utilize a metaphysical claim as a premise of the argument (a tack which Kant refuses to allow other philosophers to take).

The first of the two idealisms that Kant addresses is the Cartesian *problematic* (or skeptical) idealism. *Problematic* idealism claims (as far as Kant

is concerned) that the existence of objects external to minds is “doubtful and indemonstrable” (B 274). This is precisely the case because the existence of the Cartesian *cogito* or *res cogitans*, as our private world, is the only foundational element which is beyond doubt, whereas the existence of things or *res extensa*, as both public and external, is inaccessible to the mind. Kant insists that “the *problematic* idealism of Descartes . . . holds that there is only one empirical assertion that is indubitably certain, namely that ‘I am’” (B 274). What follows from this is that the existence of any objects in space (i.e., external to the mind of an individual) can only be causally inferred as derivative from the actual experience of objects in space. Such an inference is always open to question, and so radical skepticism concerning the existence of such objects results. It is quite clear here that the conclusion of this position involves only an epistemological claim, rather than a claim about the ontological status of the objects of experience.

It is important to note that this *problematic* idealism is not Descartes’ own position, but rather a consequence which Kant sees as immediately following from Descartes’ position (although he does not make it clear that this is the case). In the argument that follows upon his application of the methodological doubt, Descartes attempts to make the move from the certain existence of this foundational entity (the Cartesian ego) to the certain existence of a perfect and benevolent God, and again from the certain existence of this God (who, by definition, cannot be a deceiver, as deception admits of defect which no perfect being can admit) to certainty with respect to the existence of objects external to the mind. In Descartes’ mind, any question of *problematic* or skeptical idealism is answered by this rigorous deduction from the foundational, clear and distinct intuition of the *cogito* to the existence of external objects.

Kant has obviously chosen, at worst, to ignore or, at best, to reject implicitly such an argument, and to confine his attacks to the first two of Descartes’ *Meditations*, which he must consider the only points worthy of his attention in the *Meditations* as a whole. The most apparent reason for this selective Kantian reading of Descartes’ position is that his argument for the existence of external objects involves some blatantly metaphysical claims (not the least of which is the existence of God) which are necessary to secure the desired conclusion. It is precisely this sort of claim which Kant completely rejects as involving an application of the categories beyond any possible object of experience. And since Kant says that “all our knowledge begins with experience” (B 1), this God-talk is just that for

him, a groping in the dark, a misapplication of the categories, and hence, not a possible object of knowledge. Yet, if certainty with respect to the existence of God fails, then so to must certainty with respect to the existence of external objects, as the certainty of the latter depends upon and involves the certainty of the former.

The second idealism that Kant addresses is Berkelean *dogmatic* (or ontological) idealism. *Dogmatic* idealism of the Berkelean stripe (in Kant's reconstruction of it) is the position which asserts that the existence of objects external to the mind is both "false and impossible" (B 274). Kant formulated Berkeley's position as follows:

[S]pace, with all the things of which it is the inseparable condition, is something which is in itself impossible; and he therefore regards the things in space as merely imaginary entities. (B 274)

The Berkelean argument for the mind-dependence of sensible objects is what Kant is targeting here. All perception involves ideas (i.e., the idea of the thing perceived). Ideas themselves are mind dependent, and it is unthinkable to posit the existence of an idea in isolation from the mind (or in an unthinking thing). Therefore, all perception and its objects are mind dependent. Yet if perception is the perception of an idea, then it is unintelligible to posit the absolute existence of an object of perception apart from its idea, or in other words, the world is ideal. The Berkelean position, in sharp contrast with Kant's presentation of the Cartesian position, does not limit its scope to the knowability of objects, but rather involves an ontological claim about the status of the objects of experience, namely, that they are ideal.

The Kantian reconstruction of Berkeley's position obviously does not take into account the general structure of the Berkelean system. It is interesting to note, and problematic at first glance, that Kant does not even consider the possibility that Berkeley's Divine Author might fulfill the criterion of a permanence (whose importance of which will become apparent in the next section of this paper) which exists external to me as a perceiver. Berkeley is not asserting that the objects of experience are dependent upon my mind (in a solipsistic fashion), but rather they are dependent upon a mind, or specifically, God's mind. It might appear that, in virtue of this element of Berkeley's argument, he has thoroughly anticipated Kant's objections to his idealism.

It is even speculated, by many Kant scholars (Walker, 69), that Kant himself may have never read the primary texts of any of Berkeley's works, but rather that he procured his understanding of *dogmatic* idealism via second and third hand accounts as they are represented in the writings of both Hume and Beattie. Yet, in order to understand whether Kant would have accepted this notion of God as the principle of permanence, it is necessary view Kant's philosophy within a wider scope. Whether he knew of Berkeley's God or not, it is not likely Kant would have given this sort of *ad hoc* principle the dignity of an answer. Rather, he would have seen it as a purely metaphysical claim, and like all other metaphysical claims (as was the case for Descartes), it is an application of the categories to something which is not a possible object of experience. And if the certainty of the existence of objects of experience external to the perceiver is dependent upon such a metaphysical claim, then Berkeley's God is subject to a Kantian metaphysical breakdown.

II

I now must offer a more detailed analysis of Kant's specific argument against *material* idealism, as it is contained in the "Refutation of Idealism" (B 274-279). Two brief digressions, however, are necessary before beginning. The first stems from the fact that, at the beginning of the argument, Kant immediately refers the readers to the discussion of the *a priori* conditions of space and time in the "Transcendental Aesthetic" (B 69-72), in order that they might see how he has already "undermined" the grounds of the argument for Berkelean *dogmatic* idealism.

In the "Transcendental Aesthetic," Kant's dismissal of the Berkelean position is as dense as it is succinct. He states that:

When I say that the intuition of outer objects and the self-intuition of the mind alike represent the objects and the mind, in space and time, as they affect our senses, that is, as they appear, I do not mean to say that these objects are a mere *illusion* [as the Kantian reconstruction of Berkeley does]. (B 69)

Yet, Kant still wants to draw a clear distinction between representations of objects (or appearances) and things in themselves. How can Kant make good on such a claim? To begin with, Kant sees "the good Berkeley" as

insisting that space is a property of things in themselves, and that space, taken together with the objects that it contains, is impossible. If space is impossible and yet objects “seem” to be out there, then one is stuck relegating appearance to illusion. But, as Kant has argued in the “Metaphysical Exposition[s]” of space and time, space and time are not derived from the experience itself (nor are they properties of things in themselves), but rather the experience itself must be informed by the formal conditions of space and time. Kant is claiming that space and time function within human experience as a necessary condition for the possibility of the experience of objects as simultaneous or successive. As space and time are necessary conditions, and such necessary conditions cannot be derived from experience, as experience never provides anything but contingency (B 5), space and time must, therefore, be *a priori*. As space and time are the *a priori* conditions for the possibility of experience of objects, they grant “objective reality” to the objects of experience (*qua* representation) while still maintaining Kant’s distinction between appearances and things in themselves. In this way Kant claims to refute Berkeley.

Although Kant asserts that he has already disposed of the Berkelean branch of idealism in the “Transcendental Aesthetic,” the main argument of the “Refutation of Idealism” can be read as really addressing the proponents of both of these camps of idealism. Despite the fact the the argument itself, with Kant’s own admission, is focused primarily on *problematic* idealism, it would follow that, if Kant is able to prove that the existence of external objects is certain and demonstrable (against the Cartesian position that their existence is “doubtful and indemonstrable”), then it is also the case that the claim that “external objects exist” is true and, therefore, possible (against the Berkelean position that their existence is “false and impossible”). Hence, if Kant can get the best of Descartes, then he has already, of logical necessity, gotten the best of Berkeley.

The second digression is necessary to give the reader a sense of the structure of the argument against Cartesian *problematic* idealism, prior to tackling the argument itself. This will closely follow Henry Allison’s brief summary of the argument.

[T]he Cartesian skeptic conflates the consciousness of the completely indeterminate, non-individuated I of apperception, the existence of which is given—in the very act of thought, with the determinate awareness of a particular existing subject, which is given in empirical self-conscious-

ness. In short, this skepticism conflates apperception and inner sense. (Allison, 295)

It is this equation which is at the very heart of the Cartesian skeptic's argument. As is clearly evident from Allison's remark, Kant needs to maintain the clear distinction between these two. Yet, the clear and distinct *cogito* refers only to transcendental unity of apperception, and not to the empirical self. This brings into high relief the empirical self as ultimately dependent upon the objects of outer sense (and this is a point which will be further explicated in the argument itself). Therefore, (by a *reductio*) if the Cartesian skeptic is to hold out for empirical self-knowledge, then he is forced to recognize its dependence upon these external objects (Allison, 295-6).

Returning now to Kant's argument against *material* idealism, first I offer a presentation of Kant's main argument in the "Refutation of Idealism" and immediately following an analysis of the distinction between transcendental realism and transcendental idealism, in order to demonstrate that Kant sees both arguments for idealism as fundamentally based upon false metaphysical assumptions.

At the outset of the main argument of the "Refutation of Idealism," Kant clearly states his thesis as follows:

The mere, but empirically determined, consciousness of my own existence proves the existence of objects in space outside me. (B 275)

Kant defends this thesis with four premises and one conclusion. One must scrutinize each premise carefully in order to see how Kant claims to have arrived at the aforementioned conclusion.

1.) I am conscious of my own existence as determined in time.

I can verify the limits of before and after in my experience, which includes my experience of myself. This verification necessarily involves empirical self-knowledge. "[T]he . . . self is [empirically] conscious of itself [solely] as the owner of [its] sequences of representations" (Allison, 298). Any given self is aware of a particular sequence of representations of

objects of outer sense like birds, cars and its own body, as well as a sequence of representations of objects of inner sense like the flow of ideas, and as this sequence is, of necessity, informed by the a priori condition of time, so too is the consciousness of the existence of oneself determined in time, as self awareness is the awareness of a particular sequence of representations—in me.

2.) All determination in time presupposes something permanent in perception.

The first issue to be tackled here is whether this “something permanent in perception” which is presupposed is to be taken as something permanent which is perceived, or something permanent in the perceiver. The possibility of the latter of the two alternatives, as will be seen, is removed by the third premise. Therefore, that “something permanent” must be in the perception, and not in the perceiver.

Time is the measure of change, and apart from something that remains in some way the same, change is impossible. That which is permanent, must also be something which occupies space. This premise refers back to the “First Analogy of Experience” in which the principle of permanence was discussed (B 224-232). Time is itself not perceivable, and hence, this permanence must be something other than time, namely a spatial entity. It is impossible even to think of an example of a change which does not presuppose something permanent. Even in the cases of substantial change (for example, when a fire reduces a log to ashes), it is presupposed that something continues from the first state—wood—to the second state—ashes. The standard line of interpretation of this permanence is as an implicit Kantian reference to the principle of the conservation of matter. Hence, in Aristotelian terms, all change (including substantial change or a change in the very whatness of a thing) is an accidental change with respect to the permanent substantiality of matter.

3.) This permanent cannot, however, be something in me.

Assume that there is nothing outside you. Assume that all of the content of which you are conscious is nothing more than a vast array of *sensibilia*—or brute sensory data: just colors and sounds of varying intensity. There is nothing permanent in *sensibilia*: they either present them-

selves to consciousness or not. If one random grouping is not presenting itself to consciousness, another is. They have no parts to change or develop; they are just redness, or beepness, or greenness, or popness, and they are either there or they are not. Now, if this collection of brute sensory data is the entire content of my consciousness, then there is no permanence to which these *sensibilia* can be referred. And if there is no permanence, then how can I understand my own existence as determined in time (premise 1)? Time involves change. Change presupposes something permanent. But in a world composed solely of this raw sensory data, I know nothing permanent, and hence, I cannot know change. If I cannot know change, then I cannot know determination in time. If I cannot know determination in time, then I cannot know my own existence as determined in time. But, I am conscious of my own existence as determined in time; therefore, this permanent cannot be something in me.¹

The empirical self cannot be the source of permanence, as was seen in premise one. If the self were that to which this permanent refers, then there would have to be a corresponding intuition of the self which is given its structure by the category of substance. But, it will be recalled that the empirical self is itself presented through those things which are given to it in intuition, and also through the categories in terms of which the understanding is so structured to deal with what it is given in intuition, or more simply, through its contents. The self as the sum of these empirical intuitions is itself an empirical intuition which requires something permanent to underlie this continual flux amongst its representations and hence, the self as well.

4.) Hence, this permanence exists as or in something external to me as a perceiver, and not through a mere representation of something external to me. Therefore, the determination of my existence in time is only possible through the existence of actual external things which are perceived outside me.

There can only be a simple disjunction of the possibilities of the location of this permanent in perception: it either exists in the perceiver or external to the perceiver. This permanent in perception cannot exist in the perceiver (premise 3), therefore, this permanent necessarily exists as or in something external to the perceiver.

Kant obviously desires (although this desire is not unproblematic) that the existence of objects “outside me” be understood in the empirical sense. All objects of outer experience are informed by the *a priori* condition for the possibility of outer experience or space. Hence, these objects are (in the empirical sense) “outside me” as perceiver.

Three Kantian precepts will help round out the argument: (1) “all our knowledge begins with experience” (B 1); (2) knowledge of the empirical self is knowledge of its empirical intuitions or contents; and (3) “I am conscious of my existence as determined in time” (B 275). The third statement can be reduced to the claim that “I have knowledge of an empirical self.” From (2) and (3) a hypothetical syllogism can be produced, that is: “If I have knowledge of my empirical self, then I have knowledge of its intuitions or contents” and by *modus ponens* and (3) the conclusion is that “I have knowledge of its [the self’s] empirical intuitions or contents” precisely because this knowledge of empirical content is a necessary condition for the possibility of having an empirical self—period. By (1), it can be seen that empirical experience is the springboard of all knowledge, and if the knowledge of the empirical self is tied up with knowledge of its empirical contents, then prior to their introduction, the self would have been a Lockean *tabula rasa*. Therefore, apart from the perception and existence of objects of experience and more importantly the permanence through which they cohere in time, there can be no empirical self, or in other words, the consciousness of one’s existence as determined in time [inner experience] is dependent upon the consciousness of the existence of external objects [outer experience]. Again, the perception and existence of a permanence external to the self is a necessary condition for the possibility of knowledge of an empirical self.

5.) Conclusion: Consciousness of my existence as determined in time is, of necessity, the immediate consciousness of the existence of other things outside me.

Point number five succinctly reiterates what Kant claims to have proven in the “Refutation of Idealism.” Henry Allison clearly sums up this argument:

[T]he key point is that inner experience involves not merely a bare awareness of one’s representations, but also the con-

sciousness of one's existence as determined in time. The argument has shown that this consciousness requires a reference to something permanent, and that this permanent must be located in space. It therefore not only shows in a general way that the possibility of inner experience is conditioned by outer experience, it also shows that inner experience is, in fact, always correlated with outer experience. This is because the objects of both forms of experience exist in a single, universal time. (Allison, 304)

Thus, it is by clearly bringing the distinction between apperception and the empirical self (the Cartesian ego) to the fore, that Kant, by means of a reduction to the absurd, attacks the Cartesian skeptic directly, and the Berkelean dogmatist indirectly. He gives the skeptic a premise that he will accept (premise 1) and shows him what follows from it of necessity.

Kant insists that the idealists' difficulty (both Cartesian and Berkelean) with the existence of an external world is fundamentally based on a false metaphysical assumption: the assumption is that of transcendental realism. Transcendental realism is immediately characterized by the fact that it "interprets outer appearances (their reality being taken for granted) as things-in-themselves, which exist independently of us and of our sensibility, and which are therefore outside us" (A 369). The trap that such an erroneous metaphysical assumption leads one into is that of empirical idealism, and the example of George Berkeley is sufficient to make this problem clear. Kant says:

After wrongly supposing that objects of the senses, if they are to be external, must have an existence by themselves, and independently of the senses, he finds that, judged from this point of view, all our sensuous representations are inadequate to establish their reality. (A 369)

As radically distinct from (and, incidentally, the only apparent alternative to) transcendental realism, transcendental idealism avoids the aforementioned conflation of appearances and things-in-themselves by insisting that all appearances, without respect to their content are to be conceived of as "representations only, not things in themselves" (A 369). It is only appearances that are knowable; things-in-themselves do not make themselves available as material for sense experience, and hence to attempt to speak

about them is analogous to the Cartesian and Berkelean attempts to speak about God. It misapplies the categories, and does not further one's understanding of possible objects of knowledge. One must pass over things in themselves in silence. Space and time are not conditions of things in themselves, nor are they determinate self-subsisting entities, but rather, they the necessary conditions for the possibility of experience of objects (as appearances). Hence, for the transcendental idealist, the appearances of objects are empirically real, as they are precisely what they present themselves as, since in order to present themselves, they must be first constructed by reason.

III

The last issue to be considered in this paper will be a critical analysis of the "Refutation of Idealism" in order to see if such an argument can hold as it stands. This consideration will focus on three points: (1) Is Kant equivocating on what he means by the "permanent in perception"; (2) Is Kant equivocating on what he means by "objects in space outside me"; and (3) does Kant introduce an illegitimate metaphysical claim upon which his entire argument rests?

First, it is obvious that it does not make sense for Kant to call the "permanent in perception," to which he refers, phenomena. The phenomenal realm is the realm of experience, and experience provides the experienter with nothing but contingency (B 5). So, if the permanent is to be interpreted as phenomena, then the permanent itself is nothing but contingency. Such a claim, however, would be absurd, as permanence is exactly what contingency is not. But if noumena and phenomena are mutually exclusive and mutually exhaustive categories, then the permanent can be interpreted in no other way than as noumena! The difficulty here is that if the permanence is noumenal, then it cannot be in perception. Kant is guilty of an odious equivocation here which does not admit of an easy solution. The only thing that appears to be certain with respect to this permanence is that Kant wants to have his cake and eat it, too.

The second of these three issues demands a clarification on what Kant means by "objects in space outside me," and by the term "object" here Kant must be referring to the aforementioned "permanence" at least indirectly insofar as he calls both objects of perception. One thing that is clear is that Kant thinks (or at least wants to think) that these objects can only be

phenomena, and, by itself, the phrase “objects in space outside me” does not disprove this thesis. This can be the only intelligible way of interpreting the argument. For if Kant is referring to noumena, then he has broken his silence, and is making unfounded claims about things in themselves. It is on phenomenal grounds that this argument essentially rests. The phenomenal realms, however, consist solely of representations. Herein lies the equivocation. Kant explicitly states in premise four that “[the] perception of this permanent is only possible through a thing outside me and not through a mere representation of a thing outside me[.]” (B 275). Again, the phenomenal world is composed solely of appearances, which are “representations only, not things in themselves” (A 369). However, this permanent cannot be a “mere representation,” but rather must be “a thing outside me.” A thing which is not a representation, and hence not phenomenal, is (in virtue of the mutually exclusive and exhaustive disjunction) noumenal. Therefore, all the difficulties elucidated by the previous equivocation apply here as well.

It is here important to see whether or not it is possible to bring the two previous points together in order to gain greater clarity. It is clear that in the thesis for the argument, which is provided by Kant, that these objects, which exist “in space outside me,” are to be read as phenomena. Further, the next two steps that refer to these objects (premises 2 and 3) do nothing in themselves to contradict or even confuse this issue. It is, however, premise four which is the source of all this difficulty, as it is here that Kant seems to waver between the two realms. The permanence, upon which the entire argument rests, must be something which exists external to me as a perceiver, yet must not be considered a mere representation of a thing external to me. If, in the most charitable interpretation of this claim, Kant is only saying that this permanence cannot be some sort of second order representation (e.g., I have such a representation of the meal that I had several hours ago), then Kant is most certainly guilty of a fundamental ambiguity in his method of expression. He has initiated his readers into the understanding that such anti-representational language is a reference to the mind-independent things-in-themselves. If Kant has changed his use of terminology here, then some indication to this effect is in order. And even if Kant’s claim does follow the aforementioned interpretation, then the question becomes whether or not such a reading really secures the conclusion for Kant. If this permanence is, in fact, some sort of first order phenomenal representation (as distinct from the second order representation

mentioned above) then Kant is forced to accept that the permanence is actually mind dependent, as all phenomena of any order are, and as a consequence, premise three falls like a domino. It is also easier to see how Kant could fall into the trap of unintentionally missing an already difficult distinction (representation vs. things-in-themselves) than to see how he would present a premise, the logical consequence of which would contradict a previous premise (as we have already see Kant's affinity for logical consequences). And it is on these grounds that I defend the two previous arguments concerning Kant's equivocation.

Finally, one must look to see whether or not Kant is giving with one hand what he already took away with the other. Kant refuses to allow other philosophers any legitimacy with respect to the use of metaphysical claims. He rails against such nonsense by insisting that these claims apply the categories to things which are not possible objects of experience. Premise two of the "Refutation of Idealism," however, relies completely for its strength on the metaphysical claim that "nothing can change, save what is permanent in some way." Such a claim is as metaphysical as claims get. Hume would be looking for his matches to incinerate such a heretic. Hence, if metaphysics is not to be allowed, then Kant has very acutely constructed the possibility of his own 'Refutation'.

WORKS CITED

I am indebted to Dr. James Petrik (Ohio State University) for comments and criticisms on an earlier draft of this paper.

Allison, Henry. *Kant's Transcendental Idealism: An Interpretation and Defense*. New Haven: Yale University Press, 1983.

Ewing, A. C. *A Short Commentary on Kant's Critique of Pure Reason*. Chicago: University of Chicago Press, 1938.

Kant, Immanuel. *Critique of Pure Reason*. Tr. N. K. Smith. New York: St. Martin's Press, 1929.

Prichard, H. A. *Kant's Theory of Knowledge*. Oxford: Clarendon Press, 1909.

Walker, R. C. S., ed. *The Real in the Ideal*. New York: Garland Press, 1989.

¹ I am indebted to Fr. Ronald Tacelli, S.J. for pointing out this line of reasoning to me.