

## The Disappearing X

In 1786, the year before the appearance of the second ("B") edition of the Critique of Pure Reason, Kant promised that the new edition would remove the admitted "obscurity" of the first ("A") edition of 1781; the changes, however, would affect "only the manner of the exposition", not the basis of explanation.<sup>1</sup> When the B edition appeared in 1787, however, it seemed no less obscure than the A edition, and despite Kant's claims that the changes are merely cosmetic, they are often treated as substantial - without, however, any clear sense of what the matters of substance are. In what follows I examine one particular difference between the two editions, and try to show that this difference is part of a significant attempt to overcome a difficulty in the first edition.

### I.

If we look at the main text of the Critique of Pure Reason (that is, excluding the preface and introduction), we see three sections that undergo substantial revision for the B edition: the Transcendental Deduction, the Phenomena and Noumena section, and the Paralogisms. Of these, it is the two versions of the Transcendental Deduction that have drawn the most attention; but one notable difference between the A and B Deductions also distinguishes the two versions of Phenomena/Noumena and of the Paralogisms. This is the disappearance of the Transcendental Object = X. This concept plays a prominent role in the first edition versions of these three sections, but is notable by its absence in the revised versions.

The position that this difference is merely cosmetic is difficult to maintain. First, if we consider those passages in A which contain the phrase "transcendental object" [transzendentaler Gegenstand, transzendentes Objekt], with or without the "X", and examine those that were re-written for the B-edition, we find that the phrase has entirely disappeared in the re-written passages - and this despite its prominence in the passages they replaced.<sup>2</sup> Second, though some 17 passages in B retain the phrase from A, only three of these retentions are even in the neighborhood of revised passages, and only one is in one of the three substantially-revised sections mentioned above; this seems to indicate that Kant spent very little time and attention on these passages during the revision process.<sup>3</sup> And third, the "X", though it appears prominently with the phrase (and without it)<sup>4</sup> in A, appears only twice in B: once in reference to the transcendental subject<sup>5</sup>, and the other time to indicate something unknown in general, rather as

one would use "x" in a mathematical context - and even here Kant modifies A's "What is here the X...?" to "What is here the unknown something = X...?" in B.<sup>6</sup> Thus a defender of the "merely cosmetic" position on the revisions would have the task of explaining the total absence of the phrase in the rewritten parts, despite its importance in the first edition. The defender of the "significant change" position only has to justify the phrase's retention in parts of B which were revised only lightly or not at all from A; and for this it may be sufficient to note the time pressure Kant was under as the second edition was being prepared for press.

But to say that the disappearance of the Transcendental Object = X represents a significant change is not to say what its significance is. For this, we need first to show how this notion functioned in the 1781 edition, and then why this function seemed to need serious revision by 1787.

## II.

The problem of the relationship between a representation and the object it represents makes its appearance early on in the development of Kant's Transcendental philosophy; but at first only with respect to the intellectual representations (later called the Categories).<sup>7</sup> By time of the first-edition Critique in 1781, the problem had come to concern the sensible representations as well. The theoretical relation between a representation and its object has now been brought under the Copernican turn,<sup>8</sup> giving rise to a new difficulty. If the objects must "conform to the representations," they would seem to this extent dependent on the representations. But objects dependent on representations could scarcely give rise to these same representations by affecting the mind;<sup>9</sup> nor could we appeal to this affection to establish the representations' reference to the object in question.<sup>10</sup> And all this brings us to the question of just what it means in the Critical context for a representation to have "reference to an object". In the Second Analogy, Kant states the problem thus:

We have representations in us of which we can also become conscious...but...they remain just representations, i.e. inner determinations of our mind....How is it that we come to posit an object for these representations...? Objective meaning cannot consist in the reference to another representation (of what we might call the object), since then the question repeats itself....[The question under investigation is] what kind of new characteristic is given to our representations by their reference to an object, and what is the dignity that they obtain thereby....<sup>11</sup>

That is, what difference does it make to a representation to say that it is related to, refers to or represents an object? And on what basis do we attribute to it this objective reference?

The general difficulty goes back at least to Descartes: even if we grant that we can know the existence and nature of our own representations, we cannot achieve the kind of "synoptic" view often attributed to God and see the objects of our representations without the mediation of the representations themselves, together with something like "lines of reference" stretching out from our representations to their objects. We can, in a sense, "see" our representations, but not what (or whether) they represent. And we need to preserve the distinction between (1) those representations which we accept as providing a depiction of the world, and (2) those which we take to represent only fictions, wishes, plans, conjectures and the like.

It is easy enough to state this distinction, but explaining it is more difficult. We may begin by saying that the first kind represent "real objects," the second "unreal objects" (or no objects at all); but if this distinction is to be effective, we must be able to determine, for a given representation, whether it has an object, and this would seem to require once again a God's Eye perspective. The empiricist solution to this problem is to call those representations real which are the result of sensory impressions, what Kant in the 1772 letter calls "affection by the object." But it is notoriously difficult to determine whether "affection by the object" has occurred, and basing the difference between veridical and fictive representations on whether some property of affection (such as "adventitiousness") is present leads quickly to Berkeleyan idealism or Humean skepticism.

Kant too needs to distinguish those representations that are held to depict the world from those that are not. Thus, to the question: "what 'dignity' is added to a representation when we say that it refers to an object?", his answer is this:

We find that [reference to an object] does nothing more than to make the combination of the representations in a certain way necessary, and to subject them to a rule....<sup>12</sup>

Thus, not all representations will be construed as having an object, but only those whose connection with others is necessary: to say a representation refers to an object is just to say that its combination with others is necessary; and to say this combination is necessary is just to say the representation refers to an object. In sum, "having a necessary connection to other

representations" serves the function for the Critical philosophy that "arising from affection by the object" does for Empiricism: it serves to differentiate world-depicting or veridical representations from fictive ones. (And it is as a result of this new method of sorting representations into veridical and fictive, I think, that Kant is often viewed as the originator of the so-called Coherence Theory of Truth.<sup>13</sup>)

But if a representation's "having a reference to an object" just means that the representation's relations to other representations become determinate and necessary, we can establish that a representation has an object without knowing anything about the object itself at all, without, that is, knowing anything about this object that differentiates it from any other object. This consideration makes clear why Kant introduces the Transcendental Object = X the way he does in the A-Deduction:

And so it is necessary here to make understandable what we mean by the expression, an object of representations. We said above: that appearances are themselves nothing but sensible representations, which in themselves, in just the same way, must not be regarded as objects (outside the power of representation). What then do we understand when we speak of an object corresponding to the cognition, and hence differentiated from it? It is easy to see that this object must be thought only as a something in general = X, since after all we have nothing outside our cognition which we can set up against this cognition as corresponding to it.<sup>14</sup>

Kant holds that the "dignity" conveyed on representations by attributing to them objective reference - the only function of this attribution - is that it makes the representations' connections with each other necessary. As long as I do not claim that there is an object for my representations, I am free to imagine spheres which are cubical, flies which change into elephants, or halves of regions each greater than the whole. But once I say the representations have an object, I am subject to the constraints of consistency, and must make all my combinations conform to the same rules of necessity.

Kant writes:

All our representations are in fact referred by the understanding to some object or other, and since appearances are nothing but representations, the understanding refers them to a something as the object of sensible intuition: but this something is to this extent only the transcendental object.<sup>15</sup>

All representations have, as representations, their object....Appearances are the

only objects that are immediately given to us, and that which refers immediately to the object is called intuition....But these appearances are not things in themselves, but are themselves only representations, which in turn have their object, which can thus no longer be intuited by us, and hence may be called the non-empirical, i.e. transcendental object = X.<sup>16</sup>

...[This transcendental object] is one and the same for all appearances.<sup>17</sup>

It is also not difficult to see why Kant refers to this object as "transcendental." In the first place, for Kant the term "transcendental" refers to the conditions of the possibility of experience; clearly the relations among representations must be rendered necessary if experience in the Kantian sense is to be possible, and this in turn is possible only by referring them to an object. Second, "transcendental" is often used in opposition to "empirical," and whereas empirical objects can be differentiated from one another by appeal to the particular experience in question, the Transcendental Object is always the same for any appearance.

### III.

There are, however, difficulties with this Transcendental Object = X. The first is perhaps less important: it concerns the status of the Transcendental Object in the Critical system. It was noted above that one factor in favor of this object's transcendental status is that it is not empirical. But it is not only transcendental entities - space, time, and the Categories - that stand opposed to empirical objects: transcendent, or supersensible, entities do as well. The forms of intuition, space and time, are transcendental because they are conditions of the experience of empirical objects, and the pure concepts of the understanding, the Categories, are transcendental for the same reason. But the pure concepts of reason, the Ideas, differ significantly from both the forms of intuition and the Categories of the understanding. God, freedom and immortality are certainly not empirical objects, but do not have the Categories' status either: though above the empirical, they are supersensible rather than (narrowly) transcendental.<sup>18</sup>

Assuming a systematic difference between the two realms, is the Transcendental Object = X really transcendental, or is it a member of the supersensible realm? Kant seems to waver on this point. In the A-Critique's Phenomena/Noumena section, he states quite definitely that the Transcendental Object "is not the noumenon," which is in fact what we would expect from its use as a condition of possibility of experience.<sup>19</sup> Still, the references to a "supersensible ground" of the appearances, where these are construed as representations, seems to play a role in the

Dialectic similar to that of the Transcendental Object in the A-Deduction.<sup>20</sup>

But whatever the difficulties in relating these two problematic entities, the responsibility for the disappearance of the Transcendental Object = X from the passages rewritten for the B-edition is likely to be found closer to home. I think it can be found in the basically contradictory nature of the Transcendental Object's function in the A edition in general and in the A-Deduction in particular.

In the A-edition version of the Critical philosophy, the objects that Kant's predecessors had taken as things in themselves are construed as appearances, so that synthetic a priori knowledge, and hence metaphysics, could be possible. If we take Kant quite literally, however, these "objects" - the appearances - are in fact objects only in a quite diminished sense. To begin with, the appearances are only what we may call second-level representations. They serve to unite lower-level representations, such as those representing accidents, stages and parts, into empirical objects. But since ultimately they are only complexes of these lower-level representations, they need a further object to render necessary their own interrelations, and hence also those of the lower-level representations; it is for this purpose that the Transcendental Object = X is introduced.<sup>21</sup> What this means is that such objecthood as the appearances possess is borrowed in the A-edition from the Transcendental Object. But the Transcendental Object is not an empirical object at all; it seems to have more in common with the supersensible Ideas than it does with the appearances, to the extent that the appearances are taken as complexes of representations. And this means that the appearances in turn have more in common with the representations than they do with the Transcendental Object.

As a result, the A-edition view of the empirical object, as I have reconstructed it, is of an entity in serious danger, for internal reasons, of collapse. Our representations are "modifications of the mind";<sup>22</sup> but then, if their objects, the appearances, are just complexes of these representations, they too are only modifications of the mind, albeit more complex ones. And the only thing which can be construed as an object in the full sense, that is, as not needing its own object in turn, is the Transcendental Object = X - but this is a kind of ghost object, ideal construct or focus imaginarius,<sup>23</sup> and certainly not an empirical object.

When expressed this baldly, it is hard to fault Kant's first reviewers for considering his Transcendental Idealism to be a variant of Berkeleyan idealism. The basic entities are the mind

and its modifications (Berkeley's spirit and ideas, respectively). A modification of something can scarcely exist without the something modified, in the same way that on Berkeley's "esse est percipi" principle, ideas cannot exist outside of a spirit that perceives them. And it does no good to say that the lower-level modifications (representations) have second-level modifications (appearances) as their "objects" - since the appearances as second-level representations are also modifications, "reference to an object" indicates only a further relation that the various modifications, the relatively more simple ones and the relatively more complex ones alike, have to each other. The entire world becomes a complex set of relations among various modifications of the mind - among ideas in spirit, in Berkeley's terminology - and the positing of anything beyond these (such as the Transcendental Object) serves only to designate one system of these relations as privileged (that is to say, as "real"). If this is an improvement on Berkeley at all, it is at best a minor one; one might even suggest the possibility of replacing the Transcendental Object's certifying function with that of God as originator of our adventitious ideas.

There are indications, however, that already during the composition of the first-edition text, and hence prior to the first reviews, Kant had begun to recognize the difficulties involved in construing the appearance as both the object of representations and at the same time a representation itself in need of an object, and had begun tentative steps toward a solution. In the 1781 version's Second Analogy, Kant writes:

Now we can call everything, even every representation insofar as we are conscious of it, an object; but what this word is supposed to mean for appearances, not insofar as they (as representations) are objects, but only refer to an object, requires a deeper investigation. To the extent that they, only as representations, are simultaneously objects of consciousness, they are not at all different from the apprehension, i.e. the adoption into the synthesis of imagination....<sup>24</sup>

Kant then repeats the basic position:

But as soon as I raise my concept of an object up to transcendental significance, the [thing in question] is no thing in itself, but only an appearance, i.e. representation, whose transcendental object is unknown....Here, that which lies in the successive apprehension is considered as representation, but the appearance given to me, despite the fact that it is nothing more than a sum [Inbegriff] of these representations, is considered as their object....

But his attempt at a solution in the rest of this passage only deepens the obscurity:

...Appearance, in contradistinction from the representations of apprehension, can only be represented as the object different from them, if they [sc. the representations of apprehension] stand under a rule which distinguishes them from every other apprehension, and makes necessary a kind of combination of the manifold. That in the appearance which contains the condition of this necessary rule of apprehension is the object.

The object makes the representations' relations necessary, as before; but rather than say that the appearance has an object (namely, the Transcendental Object = X), Kant says that the necessity-producing object - not, nota bene, referred to here as transcendental - is something "in the appearance", presumably "in" it in the manner of a moment or an aspect, but not related to it as is elsewhere the Transcendental Object = X. My suggestion is that at this point the Transcendental Object = X begins to disappear.<sup>25</sup>

#### IV.

Thus on this reading, the central difficulty with the A-version is its identification of the empirical object, the appearance, with the representations as modifications of the mind. And there are indications in B that Kant recognized this to be the difficulty. There are numerous passages in the A-Critique identifying appearance (or object) and representation. But of the passages re-written for the B-edition, only one<sup>26</sup> contains anything like an identification of this kind: everywhere else this identification has been replaced; in places where one would expect this identification if it was Kant's intention to preserve it, it is conspicuously absent.<sup>27</sup> The elimination of this identification thus goes hand in hand with the disappearance of the Transcendental Object = X.

Furthermore, a footnote in the B-Deduction<sup>28</sup> indicates Kant's concern with the problem of the nature of representations. In the A-edition, Kant vacillates between the notion of a representation as an individual (and hence one-time, fleeting) modification of the mind, and a representation as something persisting, to which we can return.<sup>29</sup> In this note in B, however, he makes clear that every consciousness of a representation is different, so that, if I think twice of Socrates, it is not simply given that these two thoughts refer to the same object: this sameness requires that a synthesis of the two thoughts occur, which means that their sameness of reference is, like any other synthesis, an accomplishment of the subject.

What is important in the present context is that Kant here differentiates the consciousness



of something on the one hand from that of which it is the consciousness on the other. The "consciousness" is a fleeting entity existing at a particular point in time.<sup>30</sup> As the footnote points out, each one is different (if only by the difference of time-point of its occurrence), so that two "consciousnesses" can be "identified" only by being considered to be of "the same thing", and this by virtue of a synthesis accomplished by the subject. That of which the two consciousnesses are both said to be consciousness may be specified more precisely as a representation-kind; unlike the consciousnesses, this would be bound to no particular subjective time-point.<sup>31</sup> The point is that what is to be considered as a "modification of the mind" strictly speaking must be a consciousness: if we tried to construe a representation-kind, which is independent of particular subjective time-points, as a modification of the mind, there would have to be a modification of the mind existing independently of what the mind is engaged upon at a particular instant, i.e., existing when we're not conscious of it. This view would require a very complicated relationship between this representation-kind and the consciousness.<sup>32</sup> But as an empirical object, the appearance certainly cannot be different on each occasion on which it is represented, i.e. for each consciousness of it: to hold this position would be to play into the hands of the Berkeleyan interpretation, for which (or perhaps for a parody of which), objects seem to flash in and out of existence. The representation (i.e., our consciousness of it) may be different on each occasion on which it is thought, but the object it represents - the appearance - must be the same. We may synthesize the representations, but the result of this synthesis is not now a single representation stretching over both occasions: it is still two representations (consciousnesses), but now of a single object. This object, the appearance, serves the required function of making the relations among the representations necessary.

But then, what is the appearance? It is certainly not a consciousness, for it persists between occasions of my consciousness of it. It cannot be simply a representation-kind, for as a representation it would still require a further object, which is what we are trying to avoid for the appearance. And it cannot be the Transcendental Object itself, since the Transcendental Object is the same, no matter for what representation it serves as object; but we must be able to differentiate one appearance from another.

These considerations yield what I think are the basic characteristics of the appearance in the B-edition. First, it is different from a consciousness, and Kant would prefer to avoid calling

it a representation at all (though it may still be considered one in the sense that our access to it is by way of "representations" (consciousnesses) as modifications of the mind). It is more than an A-version appearance, since it is more than merely a combination of consciousnesses: not only is it the object of various present or past consciousnesses, it is the (potential) object of future ones as well, or ones that may never occur.<sup>33</sup> This makes it available as reference of consciousnesses when no consciousnesses of it are being entertained; that is, it can persist unperceived. Third, there must be differences, not only between representations, but also between their objects, so that a consciousness may refer to one object but not to another. This means that there can be more than one object, and that it can be an empirical object.

But then we have a final problem: if the function of the appearance-Transcendental Object relation in the A-edition involved an inconsistency, why did Kant not remove all traces of the Transcendental Object from B? Why are there passages retained from A which employ this expression? As noted above, time pressure may provide one answer; but I suggest it was also possible to retain the passages for the following reason. Even if the appearances are now the full-fledged objects of our representations, they all possess in common the moment of objectivity; and this could be (broadly) construed as the Transcendental Object. In the ideal, all appearances - and possible appearances - are combinable into one Nature, or one World. One could thus refer to this ideal result of combination, the World-Whole, as the Transcendental Object. It would not be an object from which the appearances borrow their objecthood, as in A; but it is true that this abstract ability to be combined consistently into one nature is a condition of the possibility of the experience of an appearance, and can thus be said to be both "transcendental," and - in the sense that a world-whole is in a (somewhat strained) sense the object of its components - also its "object."

## N O T E S

<sup>1</sup> Metaphysische Anfangsgünde der Naturwissenschaften [MAdN], Ak. IV.475f., n. (As customary, references to the Critique of Pure Reason [KdrV] cite the pagination of the first and second (A and B) original editions; all other references to Kant's work are by volume, page (and where necessary, line) of the Prussian Academy edition [Ak.] of Kant's complete works.)

<sup>2</sup> That is, the uses at A109 (Transcendental Deduction), A250ff. (Phenomena and Noumena) and A538, A361, A366, A372, A379, A390, and A393f. (Paralogisms) have no counterparts in the corresponding sections of B.

<sup>3</sup> The 17 passages are: A46/B63, A191/B236, A247/B304, A277/B333, A279/B335, A288/B344, A478/B506, A494/B522, A495/B523, A538f./B566f., A540/B568, A545/B573, A557/B585, A565/B593, A614f./B641f., A679/B707, A698/B726. A247/B304 is in the Phenomena/Noumena section, in a four-paragraph part taken over entirely into B. A46/B63 and A191/B236 are in the vicinity of modified passages - lightly modified in the first case, significantly in the second; but the modifications in this latter case (the Second Analogy) consist entirely in replacing the title and statement of the principle, and adding two introductory paragraphs, and the passage containing the retention may be regarded as transitional, as explained below.

<sup>4</sup> See especially the A-Deduction A104 and A105, where Kant explains that the "X" which is the object of our representations must be construed as transcendental, as it is explicitly called at A109.

<sup>5</sup> B404 (= A346).

<sup>6</sup> KdrV, cf. A9 and B13.

<sup>7</sup> In his well-known letter to Herz of February 21, 1772 (Ak. X.130f.), Kant asks: "On what ground rests the reference of what we call representation in us to the object? When the representation contains only the way [Art] in which the subject is affected by the object, it is easy to see how the object could be in accord with [these representations] as its effect, and how this determination of our mind could represent something, i.e. have an object.... In the Dissertation I was satisfied to express the nature of the intellectual representations [= categories] merely negatively, namely, that they were not modifications of the soul by the object. But then how else it may be possible to have a representation which refers to an object without being affected by it in any way - this I passed over in silence."

<sup>8</sup> KdrV, Bxvi. "Previously it was assumed that all our cognitions must conform to the objects, but all attempts to establish something about them through concepts which would extend our cognition came to nothing under this presupposition. Let us then try it otherwise, and see if we cannot make better progress with metaphysics' tasks by assuming that the objects must conform to our cognitions...."

<sup>9</sup> See my "Two Perspectives...", JHistPhil 32 (1994), p. 415

<sup>10</sup> Among other things, it is difficult to see how we could establish that affection has occurred without already knowing that the resulting representation refers to an object.

<sup>11</sup> KdrV, A197/B242.

<sup>12</sup> KdrV, A197/B242f.

<sup>13</sup> Cf. e.g., Proleg, Ak. IV.290.

<sup>14</sup> KdrV, A104, my emphasis.

<sup>15</sup> KdrV, A250.

<sup>16</sup> KdrV, A108f.

<sup>17</sup> KdrV, A253; cf. A109.

<sup>18</sup> Kant in fact refers to them as "transcendental ideas" (KdrV, A321/B377, *passim*), while seeming to deny that anything can be either (A296/B352f.). There seems a sense in which what is transcendent and what is transcendental is sometimes a matter of degree. This consideration, however, is beyond the scope of this paper.

<sup>19</sup> KdrV, A253. Cf., however, A366: "What kind of thing in itself (transcendental object) matter may be...."; also A393, A288/B344.

<sup>20</sup> Henry Allison (Kant's Transcendental Idealism, New Haven: Yale, 1983, p. 242-46, *passim*) takes the Transcendental Object to be identical with the noumenon. He does not, however, take the difference between the A and B editions seriously; cf. p. 246.

<sup>21</sup> KdrV, A104f.

<sup>22</sup> KdrV, A98f., A197/B242, etc.

<sup>23</sup> KdrV, A644/B672. Cf. also Ak. II.344ff. & note.

<sup>24</sup> KdrV, A189f./B234f.

<sup>25</sup> Thus on the current interpretation, this passage in the first-edition Second Analogy takes on the role of transition, and any evidence that this passage was composed after the A-Deduction and the correlated passages in the first-edition section on Phenomena and Noumena would support my thesis. Unfortunately, I have been unable so far to discover such evidence.

<sup>26</sup> KdrV, B164.

<sup>27</sup> In A, the Deduction contains some four such identifications (A105, A108f., A113, A115); Phenomena and Noumena contains one (A250); the Paralogisms contain six (A371, A372, A375n., A377, A386, A390), and the first four of these occur in that edition's version of the much-revised Refutation of Idealism, though no such identification is contained in the later versions of that argument. The only place in the revised versions of these sections where such an identification occurs is in §26 of the B-Deduction, and even here its scope seems limited to the context at hand.

<sup>28</sup> KdrV, B131 (§15).

<sup>29</sup> For the present concern it is indifferent whether the representation in question is of an individual substance, such as Socrates, or of an individual accident, such as Socrates' snub-nosed-ness: the question is whether two occasions of thinking of Socrates (or of his snub-nosed-ness, respectively) is two representations or one representation repeated.

<sup>30</sup> That is, at a particular point in the subject's stream of consciousness. I can think at 1:00 pm of the roll I ate at 7:00 am, and think of it again at 3:00 pm; these are two different consciousnesses (existing at two different times, 1:00 pm and 3:00 pm) of the same thing existing at a different time (7:00 am).

<sup>31</sup> That is, to no time-point in the subject's stream of consciousness. A representation-kind representing Socrates' death refers to 399 BC; but I can think of Socrates' death any time I want, so this representation-kind is not bound to any time-point in my mental life. (This relation would

hold whether the representation-kind represents an individual substance, an individual accident, or a general species of substance or accident-kind.)

<sup>32</sup> The A-Deduction's Threefold Synthesis, I think, was an attempt to explicate such a relationship.

<sup>33</sup> Thus as appearance a given empirical object, such as my desk, is not only the object of reference of certain consciousnesses I have had of it (of top, from side, recent past), but also of others I have not yet had and may never have (of bottom, back side, ultimate fate).