

Kant's One Self and the Appearance/Thing-in-itself Distinction

Colin Marshall, University of Melbourne

Kant's transcendental idealism hinges on a distinction between appearances and things in themselves. The debate about how to understand this distinction has largely ignored the way that Kant applies this distinction to the self. I argue that this is a mistake, and that Kant's acceptance of a single, unified self in both his theoretical and practical philosophy causes serious problems for the 'two-world' interpretation of his idealism.

The most important distinction in Kant's metaphysics is that between appearances and things-in-themselves (between phenomena and noumena),¹ as shown by the fact that it is this distinction which Kant primarily appeals to in defining 'transcendental idealism.'² Broadly speaking, Kant's readers have interpreted the distinction in two ways. One group of interpreters claims that appearances and things-in-themselves are, in some important sense, the same things. Another group of interpreters claims that appearances and things-in-themselves are not, in any important sense, the same things. The former sort of interpretations are typically called 'one-world' or 'two-aspect' interpretations, the latter, 'two-world' or 'two object' interpretations.

This terminology misleadingly suggests that all interpretations can be neatly identified as one or the other.³ There are different senses one might give to the notion of 'same things,'

¹ Also: empirical/sensible vs. intellectual. Though these are not exactly the same distinctions (cf. KrV B307 on the two senses of 'noumenon'), this will not matter for my argument (on these distinctions, see Willaschek, Marcus: "Die Mehrdeutigkeit der Unterscheidung zwischen Dingen an sich und Erscheinungen bei Kant. Zur Debatte um Zwei-Aspekte- und Zwei-Welten-Interpretationen des transzendentalen Idealismus". In *Akten des IX. Internationalen Kant-Kongresses*. Berlin 2001, 679–690).

² E.g. KrV A28/B44, A36/B52, A491/B519, Prolog, 4:293, 4:337, 4:374.

³ Anja Jauernig (Jauernig, Anja: *How to Think about Things in Themselves: An Essay on Kant's Metaphysics and Theory of Cognition*. Forthcoming) also notes that the 'two-world' title does not fit with Kant's way of individuating worlds. The distinction is not always characterized in the way I give here. Sometimes the one-world interpretation is described as that which *identifies* appearances and things-in-themselves, and the two-

however, so there is really a spectrum of interpretive views here. A one-world interpreter might ascribe to Kant the view that (a) each appearance is identical to some thing-in-itself, the view that (b) each appearance is some X-as-it-appears, where there is a thing-in-itself that is the X-as-it-is-in-itself for the same X, or the view that (c) talk of an appearance is non-metaphysical talk of an X regarded in a certain way, where one can also talk about that same X as a thing-in-itself.⁴

Two-world interpreters deny that Kant accepts any of these claims, and then give some story about what appearances and things-in-themselves are, and how they relate. This story can also take a number of forms, but typically takes all the facts about appearances to be based at some level in facts about our representations (while allowing that the facts about our representations may themselves be based at some level in facts about things-in-themselves).

In what follows, I defend the following argument for the one-world interpretation, which I call the One-Self/One-World Argument:

1. The self as it appears and the self as it is in itself are the same thing.
2. The distinction between the self as it appears and the self as it is in itself is an instance of the general appearance/thing-in-itself distinction.
3. The general appearance/thing-in-itself distinction is metaphysically uniform with respect to sameness.

world interpretation as the denial of this. This characterization is either misleading or too narrow (see following fn.). Sometimes the one-world interpretation is characterized as that which holds that appearances are *mind-independent*, but this is again too narrow (and, I argue elsewhere, too crude: Colin Marshall, “Qua-objects and Kant’s appearance/thing-in-itself distinction”). Lastly, sometimes the two-world interpretation is characterized as that which takes appearances to be constructions out of representations, and the one-world as the denial of this. This, though, leaves the one-world reading too underspecified (one could deny that appearances are constructions without accepting that they are the same things as things-in-themselves). Cf. Ralph Walker: “Kant on the Number of Worlds”. In *British Journal for the History of Philosophy* 18. 2010, 821-843, 826ff.; Dennis Schulting: “Kant’s Idealism: The Current Debate”. In *Kant’s Idealism: New Interpretations of a Controversial Doctrine*. Dordrecht 2011.

⁴ I do not claim that this is an exhaustive taxonomy. (a), (b) and (c) are sometimes treated as equivalent by Kant’s commentators, but they are logically distinct. There is related interpretive possibility here that is not easily classified: identifying appearances and things-in-themselves with different aspects of the same things (e.g. dividing up a thing’s properties into two sets, and identifying one set with an appearance and another with a thing in itself) while denying that appearances and things-in-themselves are the same things in any important sense (e.g. Langton, Rae: *Kantian Humility*. Oxford 1998, 20). Classification issues aside, such a view would also be subject to the argument that follows.

4. Therefore, appearances and things-in-themselves are the same things.

The first two premises are self-explanatory. According to the third premise, either all instances of the distinction between an appearance and the relevant thing-in-itself (or things-in-themselves) concern the same thing, or else none do. Given this and the first two premises, it follows that all appearances are the same things as the relevant things-in-themselves. Below, I defend the One-Self/One-World Argument by clarifying and supporting each premise and then considering two possible replies a two-world interpreter might make.

Though the unity of the self has informed the development of some one-world readings,⁵ the question of how well a two-world reading can make sense of the unified self is largely overlooked in the literature. It does not appear, for instance, in Lucy Allais' careful survey of objections to the two-world reading.⁶ Robert Adams and Ralph C. S. Walker note that Kant identifies the phenomenal and noumenal selves in his moral philosophy, but neither thinks this has any general consequences for the one- vs. two-world debate.⁷ Anja Jauernig, in her forthcoming book, points out that this is a consideration that a one-world interpreter "might try to spin in his favor", but thinks that a quick response suffices to dismiss it.⁸ I believe, however, that the One-Self/One-World Argument deals a significant blow to the two-world reading. The argument also implies that understanding Kant's metaphysics of the self

⁵ For instance, Prauss, Gerold: *Kant und das Problem der Dinge an sich*. Bonn, 1974; Allison, Henry: *Kant's Theory of Freedom*. Cambridge, 1990 (especially Chapter 2); Ameriks, Karl: *Kant's Theory of Mind*. Oxford, 2000 (especially Chapters 6 and 7). In a related vein, Desmond Hogan has recently argued that Kant's views on freedom need to be given a central place in understanding even his basic epistemological claims about things in themselves (Hogan, Desmond: "Noumenal affection." In *Philosophical Review* 118, 2009, 501-532.).

⁶ Lucy Allais: "Kant's One World: Interpreting 'Transcendental Idealism'". In *British Journal for the History of Philosophy*, 12. 2004. 655–684, 660-665.

⁷"the moral argument only allows such a [phenomenal-noumenal] identification where rational beings are concerned... clearly there can be no prospect of extending the moral argument to cover ordinary physical things" (Walker, "Kant on the number of worlds", 837). See also Robert Adams: "Things In Themselves". In *Philosophy and Phenomenological Research* 57, 1997, 801-825, 821ff.; Karl Ameriks: *Interpreting Kant's Critiques*. Oxford 2003, 84; and Wagner, Hans: "Kants affirmative Metaphysik von Dingen an sich". In *Zu Kants Kritischer Philosophie*. Würzburg 2008, 76-77. Interpreters who are aware of the potential force of this criticism of two-world readings include Allen Wood, Tobias Rosefeldt, and Henry Allison (Wood, Allen: *Kant*. Oxford. 2005, 74; Rosefeldt, Tobias: "Dinge an sich und sekundäre Qualitäten". In *Kant in der Gegenwart*, ed. By J. Stolzenburg. Berlin and New York 2007, 167–209, 170, Allison, Henry: *Essays on Kant*. Oxford 2012, 80). None of these interpreters fully develop the criticism, however.

⁸ Jauernig, *How to Think about Things in Themselves*.

is more directly relevant to interpreting his transcendental idealism than many of his readers have assumed.

One caveat: my argument rests on a fairly weak assumption of charity. I do not assume that Kant is perfectly consistent. All my argument assumes is that there is broad consistency between those core pieces of Kant's philosophy whose interrelations he explicitly and repeatedly discusses. Kant may well have changed his mind about many aspects of his idealism, but the premises of the One-Self/One-World Argument rest on tenets of his philosophy that we find throughout his mature works.

1. The self as it appears and the self as it is in itself are the same thing

Kant ascribes both phenomenal and noumenal aspects to us. In the relevant passages, he uses several notions for 'us': the I, the self, the subject, the mind, the person and the human. Though there are important differences between these, I use 'self' as a blanket term in what follows.⁹

How many selves are there, then? The common-sense view is that each of us has exactly one self. All of Kant's major philosophical predecessors accepted this, though they made sense of it in different ways. If Kant had rejected this 'single self' view, we would expect him to do so explicitly. Yet in the *overwhelming* majority of places where Kant discusses the self, he talks about a single thing.¹⁰ Only on rare occasions does he use qualified terminology (e.g. explicitly restricting a claim to the self as it appears or as it is in itself), and his talk of distinct selves is rarer still. Despite this, many of Kant's commentators imply that he consistently distinguishes the phenomenal/empirical self from the noumenal/intelligible self. James Van Cleve, for instance, states:

⁹ Kant's notion of the self involves more detail than I can properly discuss here. I discuss the relevant distinctions in Colin Marshall: "Kant's metaphysics of the self". In *Philosophers' Imprint* 10, 2010, 1-21, 2. Perhaps the most obvious distinction concerns whether just the appearances of inner sense are included (as is true of the mind), or also those of outer sense (as is true of the embodied human being). On the (at least possible) unity of the thinking self with the body Kant says: "I can [...] assume that [...] the very same thing that is called a body in one relation would at the same time be a thinking being in another" ("kann ich [...] annehmen [...] [daß] was in einer Beziehung körperlich heißt, in einer andern zugleich ein denkend Wesen sein") (KrV, A359).

¹⁰ This is also true of his pre-critical work. See Klemme, Heiner: *Kants Philosophie des Subjects*. Hamburg. 1995, 38-138 and Wuerth, Julian: "Kant's Immediatism, Pre-Critique". In *Journal of the History of Philosophy* 44, 2006, 489-532.

Kant speaks of three different selves: the empirical self, the transcendental self, and the noumenal self. The empirical self is the self as encountered in introspection... The transcendental self is the subject of all the states just mentioned: it is the thinker of our thoughts, the haver of our experiences, the willer of our actions, and perhaps also the agent of the various activities of synthesis Kant talks about... the noumenal self is the ‘self in itself’ – the real self or the self as it really is.¹¹

This is misleading. Kant almost never talks of distinct selves.¹² In fact, the phrases 'empirical self,' 'transcendental self,' and 'noumenal self' appear nowhere in any of the *Critiques*. One could respond that Kant only makes the relevant distinctions (perhaps in other terms) when he is being ‘serious,’ and that his usual talk of a single self is merely for presentational purposes. But the text does not support this response. For instance, in the dense discussion of inner sense and apperception in §25 of the B Deduction, Kant writes that “[i]n the transcendental synthesis of the manifold of representations in general... I am conscious of myself not as I appear to myself, nor as I am in myself, but only that I am” (B157).¹³ This is surely a serious discussion of the self, and yet Kant writes of a single self that is intuited, thought, appears, and is in itself. To be sure, whether we should take Kant at his word hinges

¹¹ James Van Cleve: *Problems from Kant*. Oxford. 1999, 182. Van Cleve goes on to argue (quite plausibly) that the noumenal and transcendental selves are the same (Van Cleve, *ibid.* 182-85), but maintains a radical distinction between these and the empirical self. See also Schulting, “Kant’s Idealism”, 168-170. This assumption is reflected in what is described as the issue of ‘Kant’s third self.’ See Kitcher, Patricia: “Kant’s Real Self”. In *Self and Nature in Kant’s Philosophy*. Ithaca. 1984, 113-147, 113 and Zöller, Gunter: “Main Developments in Recent Scholarship on the *Critique of Pure Reason*”. In *Philosophy and Phenomenological Research* 53. 1993, 445-66.

¹² At GMS 4:457, Kant writes of the human having his ‘proper self’ in the intelligible world, but the context makes it clear that the human being is one thing, also having a phenomenal side. KprV 5:161 talks of the ‘invisible self,’ but this is an invisible self that is said to belong to someone who also has an animal nature. Kant does (as we will see below) talk of us having distinct phenomenal and noumenal *characters*, but he consistently ascribes those characters to a single entity (the self). In FM/Lose Blätter 20:338, Kant does talk of the “empirical I” (“das empirische Ich”) vs. the “rational I” (“das rationale Ich”) and of the “logical I” (“das logische Ich”) vs. the “physical self” (“das physische selbst”), but simultaneously appeals to the identity of the self.

¹³KrV, B157: “Dagegen bin ich mir meiner selbst in der transcendentalen Synthesis des Mannigfaltigen der Vorstellungen überhaupt [...] bewußt, nicht wie ich mir erscheine, noch wie ich an mir selbst bin, sondern nur daß ich bin”. See also KprV 5:114, 5:97 and 5:105.

on whether we can make philosophical sense of there being a single self that has both phenomenal and noumenal aspects, but it is plausible that this can be done.¹⁴ I say a bit more at the end of this section about the textual constraints on how to understand the singleness of the self.

It is possible that Van Cleve and others have been misled by Kant's distinctions between types of *self-representation*. Kant distinguishes empirical apperception and inner sense from pure apperception, for instance (e.g. KrV A107, B132), but this no more implies that there is more than one type of self than the fact that there are empirical and pure concepts of objects by itself implies that there is more than one type of object.¹⁵

The textual grounds for the first premise are therefore extremely strong. We find more systematic grounds in the central themes of Kant's practical philosophy.¹⁶ Two strands of the practical philosophy are directly relevant: our necessitation by objective practical laws (Kant's view of duty), and our actions' having both empirical and non-empirical causes (Kant's compatibilism). I focus on the issue of practical necessitation here, returning to compatibilism in the next section.¹⁷

According to Kant, our relation to objective practical laws is halfway between the relation a (non-empirical) divine being has to them and the relation a non-rational animal has. A divine being's actions are directly determined by its representation of these laws (cf. GMS 4:414). A non-rational animal has no representation of these laws, and acts merely on the basis of empirical incentives (cf. KprV 5:61). The human will is subject to both rational and

¹⁴ For one approach, see my "Kant's metaphysics of the self".

¹⁵ See Klemme, *Kants Philosophie des Subjekts*, 214ff. At 7:142 in the *Anthropology*, Kant claims that "I as a thinking being am one and the same subject with myself as a sensing being" (Anth, 7:142: "Ich, als denkendes Wesen, bin zwar mit Mir, als Sinnenwesen, ein und dasselbe Subjekt".) and in the next paragraph complains about psychologists who confuse inner sense with apperception. So Kant clearly distinguished the question of how many types of self-representation there are from the question of how many selves there are.

¹⁶ There are also systematic grounds in the theoretical works. The unity of the sensible and non-sensible sides of the self seems to be at the core of the Transcendental Deduction (e.g. KrV, A107, B140), and the Third Analogy hinges on the experiencing self entering into causal relations with empirical objects (KrV, A213/B260).

¹⁷ Walker, "Kant on the number of worlds", 836 argues that some close tie between the phenomenal and noumenal selves is needed for cross-temporal ascriptions of moral responsibility to human beings. I think this argument is basically right, though a two-world interpreter who was willing to claim knowledge of noumenal identities could respond that Kant's view was just that a certain noumenal self accompanied the human body without being the same thing as it.

non-rational forces, however.¹⁸ Kant describes this view in grandiose terms at the end of the second *Critique*:

Two things fill the mind with ever new and increasing admiration and reverence... the starry heavens above me and the moral law within me [...] The first begins from the place I occupy in the external world of sense [...] The second begins from my invisible self, my personality, and presents me in a world which has true infinity but which can be discovered only by the understanding, and I cognize that my connection with that world (and thereby with all those visible worlds as well) is not merely contingent, as in the first case, but universal and necessary. The first [...] annihilates, as it were, my importance as an animal creature [...] The second, on the contrary, infinitely raises my worth as an intelligence by my personality, in which the moral law reveals to me a life independent of animality and even of the whole sensible world¹⁹

As Kant describes it, then, our distinctive status hinges on our being both phenomenal and noumenal. So, if we take him at his word, it is crucial to Kant's central philosophical concerns that the self is a single thing with both noumenal and phenomenal sides. Matters would be different if there were two selves, a noumenal and a phenomenal one. That would seem to be more like a case of an angel pushing around an animal. In the *Metaphysics of Morals*, Kant is clear that the singleness of the self is crucial to understanding our duties:

The subject that is bound, as well as the subject that binds, is always the human being only; and though we may, in a theoretical respect, distinguish soul and body from each other, as natural characteristics of a human being, we may not think of them as different

¹⁸ See, e.g., GMS, 4:413, KprV, 5:20, MS, 6:387, 6:420

¹⁹KprV, 5:161-62: "Zwei Dinge erfüllen das Gemüth mit immer neuer und zunehmender Bewunderung und Ehrfurcht [...] der bestirnte Himmel über mir und das moralische Gesetz in mir. [...] Das erste fängt von dem Platze an, den ich in der äußern Sinnenwelt einnehme [...] Das zweite fängt von meinem unsichtbaren Selbst, meiner Persönlichkeit, an und stellt mich in einer Welt dar, die wahre Unendlichkeit hat, aber nur dem Verstande spürbar ist, und mit welcher (dadurch aber auch zugleich mit allen jenen sichtbaren Welten) ich mich nicht wie dort in blos zufälliger, sondern allgemeiner und nothwendiger Verknüpfung erkenne. Der erstere Anblick [...] vernichtet gleichsam meine Wichtigkeit, als eines thierischen Geschöpfs [...] Der zweite erhebt dagegen meinen Werth, als einer Intelligenz, unendlich durch meine Persönlichkeit, in welcher das moralische Gesetz mir ein [...] von der ganzen Sinnenwelt unabhängiges Leben offenbart". See also KU, 5:403.

substances putting him under obligation, so as to justify a division of duties to the body and duties to the soul ²⁰

It would be a serious cost to an interpretation it could not make sense of this tenet of Kant's practical philosophy.²¹ As I emphasized earlier, moreover, such an interpretation would be forced to reject the natural reading of vast amounts of text. In slogan form: dividing the self is deeply un-Kantian.²²

In what follows, it will help to have a clear sense of how Kant describes the dual character and singleness of the self, though I continue to remain neutral on how that singleness should be understood. Two of the clearest formulations appear in the resolution to the Third Antinomy:

for a subject of the world of sense we would have first an empirical character [...] one would also have to allow this subject an intelligible character, through which it is indeed the cause of those actions as appearances, but which does not stand under any conditions of sensibility and is not itself appearance. The first one could call the character of such a thing in appearance, the second its character as a thing in itself.²³

²⁰MS, 6:418-19: "Das verpflichtete sowohl als das verpflichtende Subject ist immer nur der Mensch, und wenn es uns in theoretischer Rücksicht gleich erlaubt ist im Menschen Seele und Körper als Naturbeschaffenheiten des Menschen von einander zu unterscheiden, so ist es doch nicht erlaubt sie als verschiedene den Menschen verpflichtende Substanzen zu denken, um zur Eintheilung in Pflichten gegen den Körper und gegen die Seele berechtigt zu sein".

²¹ For a helpful discussion, see Chapter 3 of Marcus Willaschek: *Praktische Vernunft. Handlungstheorie und Moralbegründung bei Kant*. Stuttgart/Weimar. 1992.

²² Another interpretive route is to deny that the self is a thing at all (e.g. P. F. Strawson: *The Bounds of Sense*. London. 1966, 166), in which case there would be no issue of unity. This interpretation is hard to motivate, though (cf. Robert Howell" "Kant, the 'I think,' and self-awareness". In *Kant's Legacy*, ed. by P. Cicovacki. Rochester. 2001, 117-152, 121, and my "Kant's metaphysics of the self", 8-9.

²³KrV, A539/B567: "da würden wir an einem Subjecte der Sinnenwelt erstlich einen empirischen Charakter haben [...]Zweitens würde man ihm noch einen intelligibelen Charakter einräumen müssen, dadurch es zwar die Ursache jener Handlungen als Erscheinungen ist, der aber selbst unter keinen Bedingungen der Sinnlichkeit steht und selbst nicht Erscheinung ist. Man könnte auch den ersteren den Charakter eines solchen Dinges in der Erscheinung, den zweiten den Charakter des Dinges an sich selbst nennen."

the human being [...] is in one part phenomenon, but in another part [understanding and reason] he is a merely intelligible object²⁴

These passages describe the issue in metaphysical terms. In the first passage, Kant claims that the subject has two characters, an intelligible one and an empirical one. This fits reasonably well with the formulation in the second passage, where he talks of the human having *parts*. I refer to these below as 'two-part' formulations of the singleness of the dual-character self.

Kant also ascribes the dual character to a single faculty of the self (the will):

just the same will is thought of in the appearance (in visible actions) as necessarily subject to the law of nature and to this extent not free, while yet on the other hand it is thought of as belonging to a thing in itself as not subject to that law, and hence as free²⁵

one can consider the causality of this being in two aspects, as intelligible in its action as a thing in itself, and as sensible in the effects of that action as an appearance in the world of sense.²⁶

I do not think much turns on whether one privileges talk of duality of the *self* or talk of the duality of the *will*; if the will can be both phenomenal and noumenal without losing its unity, then the unified self can likewise be phenomenal and noumenal. The feature of these formulations that does merit attention is their use of representational terms: the will is 'thought of as' or 'considered as' phenomenal or as noumenal. Different one-world interpretations take this use of representational terms more or less metaphysically. Without taking a stand on that question, I refer to this latter sort of formulations as 'two-aspect' formulations of the singleness of the self.

²⁴KrV, A546/B574: “der Mensch [...] ist sich selbst freilich eines Theils Phänomen, anderen Theils aber, nämlich in Ansehung gewisser Vermögen, ein bloß intelligibeler Gegenstand”.

²⁵ KrV, Bxxvii-xxviii: “so wird eben derselbe Wille in der Erscheinung (den sichtbaren Handlungen) als dem Naturgesetze nothwendig gemäß und so fern nicht frei und doch andererseits als einem Dinge an sich selbst angehörig jenem nicht unterworfen, mithin als frei gedacht”.

²⁶ KrV, A538/B566: “so kann man die Causalität dieses Wesens auf zwei Seiten betrachten, als intelligibel nach ihrer Handlung als eines Dinges an sich selbst, und als sensibel nach den Wirkungen derselben als einer Erscheinung in der Sinnenwelt.” In the Vigilantius lectures, Kant says that our power of choice is “*partim sensualis, partim intellectualis*” (V-MS/Vigil, 29:1016). In addition to two-aspect talk (e.g. KprV, 5:114), the second *Critique* also talks of one and the same being *belonging* to two worlds (see KprV, 5:105).

Of course, one could doubt whether the above passages should be read literally. In fact, the following sections show that a two-world interpreter has reason to find some non-literal readings. The onus, though, would be for the two-world interpreter to defend a non-literal reading and to show that it does not face an analogue of the following problems.

2. The distinction between the self as it appears and the self as it is in itself is an instance of the general distinction between appearances and things-in-themselves

A two-world interpreter might respond at this point by insisting that the appearance/thing-in-itself distinction just has a different meaning when applied to the self than when applied to other entities. I suspect that many interpreters have been attracted by this line of thought. It is important to recognize, then, that the text does not support it. There is both a general and a specific textual argument for the claim that Kant takes himself to be univocally applying the appearance/thing-in-itself distinction to the self and to other entities.

The general argument is this: throughout all of his mature writings, Kant uses the same terminology in discussing the phenomenal/noumenal aspects of the self and in discussing the phenomenal/noumenal aspects of other entities. Were it the case that he really has in mind two different distinctions, we would expect him to clarify this at some point: Kant rarely shies away from making distinctions among distinctions.

The specific textual argument stems from passages where Kant directly applies the general distinction to the self. Such direct application suggests that he thinks there is nothing distinctive about the meaning of the distinction in the case of the self.

This direct application is conspicuous in Kant's discussions of his compatibilism. Here are three relevant passages. The first appears in the B Preface:

if the critique has not erred in teaching that the object should be taken in a twofold meaning, namely as appearance or as thing in itself; if its deduction of the pure concepts of the understanding is correct, and hence the principle of causality applies only to things taken in the first sense, namely insofar as they are objects of experience, while things in the second meaning are not subject to it; then *just the same will* is thought of in the appearance (in visible actions) as necessarily subject to the law of nature and to this extent not free, while yet on the other hand it is thought of as belonging to a thing in

itself as not subject to that law, and hence as free, without any contradiction hereby occurring²⁷

Kant begins with the general appearance/thing-in-itself distinction, and directly moves to the case of the single will. Indeed, as the B Preface paints it, part of Kant's motivation for generally distinguishing appearances from things-in-themselves is to make room for his compatibilism. If that is right, then it is crucial that the application of the distinction be univocal.

We see the same pattern in the resolution to the Third Antinomy. At the beginning of the section, “The possibility of causality through freedom united with the universal law of natural necessity,” Kant writes:

I call intelligible that in an object of sense which is not itself appearance. Accordingly, if that which must be regarded as appearance in the world of sense has in itself a faculty which is not an object of intuition through which it can be the cause of appearances, then one can consider the causality of this being in two aspects, as intelligible in its action as a thing in itself, and as sensible in the effects of that action as an appearance in the world of sense [...] [this] does not contradict any of the concepts we have to form of appearance and of a possible experience. For since these appearances, because they are not things in themselves, must be grounded in a transcendental object [...] nothing hinders us from ascribing to this transcendental object, apart from the property through which it appears, another causality that is not appearance²⁸

²⁷KrV, Bxxvii-xxviii “Wenn aber die Kritik nicht geirrt hat, da sie das Object in zweierlei Bedeutung nehmen lehrt, nämlich als Erscheinung oder als Ding an sich selbst; wenn die Deduction ihrer Verstandesbegriffe richtig ist [...] so wird eben derselbe Wille in der Erscheinung [...] als dem Naturgesetze nothwendig gemäß und so fern nicht frei und doch andererseits als einem Dinge an sich selbst angehörig jenem nicht unterworfen, mithin als frei gedacht, ohne daß hierbei ein Widerspruch vorgeht” (my emphasis). Another telling passage comes just before at KrV Bxxvii.

²⁸KrV, A538-39/B566-67: “Ich nenne dasjenige an einem Gegenstande der Sinne, was selbst nicht Erscheinung ist, intelligibel. Wenn demnach dasjenige, was in der Sinnenwelt als Erscheinung angesehen werden muß, an sich selbst auch ein Vermögen hat, welches kein Gegenstand der sinnlichen Anschauung ist [...] so kann man die Causalität dieses Wesens auf zwei Seiten betrachten, als intelligibel nach ihrer Handlung als eines Dinges an sich selbst, und als sensibel nach den Wirkungen derselben als einer Erscheinung in der Sinnenwelt [...] [Diese] widerspricht keinem von den Begriffen, die wir uns von Erscheinungen und von einer möglichen Erfahrung zu machen haben. Denn da diesen, weil sie an sich keine Dinge sind, ein transscendentaler Gegenstand zum Grunde liegen muß, der sie als bloße

So here again Kant uses his general distinctions (intelligible vs. sensible, appearance vs. thing-in-itself) to open a careful discussion of compatibilism that was preserved in both editions of the *Critique*. That compatibilism is consistently described as applying to a single self.

The third illustrative passage comes from the third part of the *Groundwork*:

all representations which come to us involuntarily (as do those of the senses) enable us to cognize objects only as they affect us and we remain ignorant of what they may be in themselves... As soon as this distinction has once been made [...] then it follows of itself that we must admit and assume behind appearances something else that is not appearance, namely things in themselves [...] This must yield a distinction, although a crude one, between a world of sense and the world of understanding, the first of which can be very different according to the difference of sensibility in various observers of the world while the second, which is its basis, always remains the same. Even as to himself, the human being cannot claim to cognize what he is in himself through the cognizance he has by inner sensation ²⁹

Unlike the second *Critique*, the *Groundwork* avoids explicit appeals to the doctrines of the first *Critique*. So if Kant took the appearance/thing-in-itself distinction to apply differently to the self than to other entities, we would expect him to only introduce it with respect to the self here and to stay silent on the general nature of objects. Instead, he provides the general distinction, and then considers the self as a special case.

In fact, it is plausible that the instance where Kant most cared about the appearance/thing-in-itself distinction was with the self, since he sometimes suggests that the

Vorstellungen bestimmt, so hindert nichts, daß wir diesem transcendentalen Gegenstände [...] nicht auch eine Causalität beilegen sollten, die nicht Erscheinung ist”.

²⁹GMS, 4:450-51: “Es ist eine Bemerkung, welche anzustellen eben kein subtiles Nachdenken erfordert wird [...] daß alle Vorstellungen, die uns ohne unsere Willkür kommen (wie die der Sinne), uns die Gegenstände nicht anders zu erkennen geben, als sie uns afficiren, wobei, was sie an sich sein mögen, uns unbekannt bleibt [...] Sobald dieser Unterschied [...] einmal gemacht ist, so folgt von selbst, daß man hinter den Erscheinungen doch noch etwas anderes, was nicht Erscheinung ist, nämlich die Dinge an sich, einräumen und annehmen müsse [...] Dieses muß eine, obzwar rohe, Unterscheidung einer Sinnenwelt von der Verstandeswelt abgeben [...] Sogar sich selbst und zwar nach der Kenntniß, die der Mensch durch innere Empfindung von sich hat, darf er sich nicht anmaßen zu erkennen, wie er an sich selbst sei.” see also Prol, 4:337.

core claims of his moral philosophy are of greater importance than those of his theoretical philosophy.³⁰ If that is right, then, contrary to the usual pattern of discussion, the self should be the starting place for understanding the distinction.

Regardless of whether we accept this last claim, it should be clear that if we accept that the appearance/thing-in-itself distinction is univocal, and accept the claim of the first premise of the One-Self/One-World Argument (namely, that the phenomenal and noumenal selves are the same thing), then there is at least a *prima facie* case for accepting the core claim of the one-world interpretation.

3. The general appearance/thing-in-itself distinction is metaphysically uniform with respect to sameness.

Even granting the first two premises of the One-Self/One-World Argument, it might still seem that a two-world interpreter could insist that the general appearance/thing-in-itself distinction can (without equivocation) be applied both to *relata* that are the same thing (in some relevant sense) or to *relata* that are different things. To deny the third premise in this way, the two-world interpreter might appeal to an analogy: there is a general distinction between prime ministers and historians, but the fact that Winston Churchill the prime minister was the same thing (person) as Winston Churchill the historian does not show that Tony Blair is the same person as some historian. Though univocal, the prime minister/historian distinction is not metaphysically uniform with respect to sameness. If the appearance/thing-in-itself distinction were structurally similar, then it could be the case that the self alone was both appearance and thing-in-itself.

As before, there are both general and specific reasons for accepting the third premise. The general reason is as follows. The appearance/thing-in-itself distinction is a *fundamental metaphysical* distinction in Kant's philosophy. Generally, if a metaphysical distinction is fundamental, then it is metaphysically uniform with respect to sameness (e.g. nothing is both abstract and concrete, nothing is both universal and particular). Part of the reason for this is that non-uniform distinctions are typically explained in terms of more fundamental distinctions, so we should expect uniformity at the fundamental level. All this implies that the appearance/thing-in-itself distinction is uniform with respect to sameness.

³⁰ E.g. KrV, Bxxix-xxx.

The more specific reasons for accepting the third premise are the focus of the rest of this section. Returning to the above example: if one spells out the prime minister/historian distinction, it becomes clear why it can apply in the way that it does: being a historian is about studying history, being a prime minister is about occupying a certain role in a government. These properties are compatible (unlike the properties involved in, e.g., the prime number/prime minister distinction) but have no direct entailment relations (unlike the properties involved in, e.g., the prime number/natural number distinction). So the two-world interpreter needs to show that the appearance/thing-in-itself distinction is like the prime minister/historian distinction, by spelling out the notions of appearances and things-in-themselves in such a way that the relevant properties are compatible without one entailing the other (in particular, without ‘is an appearance’ entailing ‘is a thing in itself’). The two-world interpreter needs compatibility in order to accept the unity of the self, and needs non-entailment in order to distinguish her view from the one-world interpretation. My main contention in this section is that the two-world interpreter cannot spell out the distinction in the needed way.

These concerns bear differently on different versions of the two-world interpretation. Instead of trying to survey all versions of this interpretation, I focus on what is probably the most influential version in the current literature, and certainly one of the most clear: that of James Van Cleve. By the end of the section, it should be obvious how my challenge generalizes to other versions of the two-world reading.

Van Cleve motivates his two-world reading by appealing to a metaphysical principle: “An object can depend on us for its *Sosein* (its being the way it is) only if it also depends on us for its *Sein* (its being, period).”³¹ I take it that what Van Cleve means is that if an object has a mind-dependent property, then its existence is mind-dependent. Mind-dependent existence would mean that the object’s existence (the entire object, not just some aspects or parts of it) is *metaphysically downstream* from that of the mind. Van Cleve takes appearances to have mind-dependent properties. Given the principle, this entails that their existence is mind-dependent. Since the existence of things-in-themselves is mind-independent,³² and the same thing cannot have both a mind-dependent and a mind-independent existence, it would then follow that appearances and things-in-themselves are not the same things.

³¹ Van Cleve, *Problems from Kant*, 5

³² Some of Kant’s readers have denied that things-in-themselves are mind-independent (e.g. Pogge, Thomas: “Erscheinungen und Dinge an sich”. In *Zeitschrift für philosophische Forschung* 45. 1991, 489-510), but this is a minority view.

However, consider what happens if we apply this principle to the self as appearance (what Van Cleve calls 'the empirical self'). Being an appearance, this self has mind-dependent properties (e.g. temporality). As we have seen, there is only a single self. So Van Cleve's principle entails that the self has a mind-dependent existence, so that the self's existence is metaphysically downstream from that of the mind. But for Kant, the self just is the mind, and the same thing cannot be metaphysically downstream from itself. Given the singleness of the self, Van Cleve's principle thus leads to absurdity.³³ If something like this principle has motivated other two-world interpretations, the singleness of the self undermines their motivation as well.

The central claim of Van Cleve's interpretation does not require the *Sosein-Sein* principle, however. That claim is that appearances are *virtual objects*. According to Van Cleve, to say that a "virtual object of a certain sort (e.g., a patch of red) exists is shorthand for saying that a certain kind of representation occurs."³⁴ Virtual objects have an ontological status quite different from non-virtual objects: "We would not quantify over virtual objects in an ontologically perspicuous language."³⁵ On this interpretation, it is clear that appearances are in no important sense the same things as things-in-themselves, since (assuming we could refer to them) we surely would quantify over things-in-themselves in an ontologically perspicuous language. This would show that being an appearance does not entail being a thing-in-itself (the relevant non-entailment).

Yet this way of understanding the appearances/things-in-themselves distinction makes it hard to see how there could be an instance in which a single thing was both phenomenal

³³ There are direct grounds for rejecting the principle, if 'being, period' is taken strongly. A statue is dependent on a sculptor for its *Sosein*. Yet the statue's clay clearly doesn't depend on the sculptor. This suggests that objects can have dependent properties while having independent constituents. I think this is in fact the right way to think about objects for Kant (and argue for it in "Qua-objects and Kant's appearance/thing-in-itself distinction). Less interestingly, there are challenges to Van Cleve's principle if 'the way it is' includes all properties; surely an object can depend on us for its name (e.g. 'Mount Taylor') without depending on us for its existence. This challenge could be met by restricting the principle to a certain set of important properties (Mike Raven suggested to me: essential properties), but since all the important properties of the phenomenal self depend on the noumenal self, such a move would not help with the objection that follows.

³⁴ Van Cleve, *Problems from Kant*, 8-9.

³⁵ Van Cleve, *Problems from Kant*, 11.

and noumenal (i.e. the compatibility of these properties), as the self surely is.³⁶ This is easiest to see with the two-part formulations of the self. If the self has a phenomenal part and a noumenal part, then phenomenal and noumenal things must be able to stand in some kind of *composition* relation. But it is implausible that any composition relation can hold between things we would quantify over and things we would not. Even contemporary metaphysicians who favor unrestricted views of composition (which Kant certainly did not³⁷) tend to see composition as only holding between entities of the same general category, and none that I know of accept composition relations between entities we ultimately would quantify over and those we would not.³⁸ Therefore, the more metaphysically dissimilar the phenomenal and noumenal are, the harder it becomes to see how one thing could have phenomenal and noumenal parts. Van Cleve makes them radically dissimilar, so the problem is especially acute for his reading.

Kant does allow that one type of composition relation can hold between parts of somewhat different natures: what he calls the combination of homogeneous vs. non-homogeneous manifolds (cf. KrV B201-202n.). But his examples of non-homogeneous combination (what he calls '*nexus*') are the traditional ones: substance/accident, cause/effect. Nothing suggests that he would accept combinations of virtual objects with non-virtual objects. One might think that Van Cleve could respond by subsuming the noumenal-self/phenomenal-self relation to either the substance/accident relation or to the cause/effect relation, but this would quickly lead to a general one-world view. For if the phenomenal self stands in one of these relations to the noumenal self, then surely every appearances stands in that relation to something noumenal. Any principle a committed two-world interpreter can find to unify the phenomenal and noumenal sides of the self must be a principle that does not

³⁶ Van Cleve explicitly applies the 'virtual object' reading to the empirical self (Van Cleve, *Problems from Kant*, 186). He himself seems to see no reason to connect the empirical and noumenal selves (ibid., 300). So Van Cleve is consistent in his interpretation, but does not appreciate the cost of denying the singleness of the self.

³⁷ Cf. Kant's insistence on the conditions of unity of representations (e.g. B129-30). The weakest example of combination Kant recognizes is that of "the two triangles into which a square is divided by the diagonal" (B201n.). For a helpful discussion, see David Bell: "Some Kantian thoughts on propositional unity". In *Proceedings of the Aristotelian Society, Supp. Volumes 75*, 2001, 1-16.

³⁸ David Lewis takes an extremely permissive view, allowing mereological relations between concrete particulars and abstracta (Lewis, David: *On the Plurality of Worlds*. Oxford. 1986, 212. Even the extremely permissive view, however, falls short of what Van Cleve would need. For a relevant discussion, see Hudson, Hud: "Confining composition". In *Journal of Philosophy* 103, 2006, 631-651. The point applies even more strongly to identity relations. Cf. Walker, "Kant on the number of worlds", 825.

lead to a general unification of the phenomenal and the noumenal. Otherwise, the result would be a one-world interpretation.

The objection in the previous two paragraphs relied on the two-part formulations of the self's unity. The same objection applies if we consider the two-aspect formulations. It might seem that the 'considered as' and 'thought as' locutions in these formulations provide Van Cleve with a way of uniting the self. For a single entity can indeed be *considered* in very different ways. If we, say, construe Van Cleve's 'virtual objects' along the lines of fictional entities, then we could say that the phenomenal self is something like a fictional version of the noumenal self. There would be nothing incoherent about that: *The Social Network* may or may not be largely fictional, but there is no doubt that its main character is (in some important sense) the same person as the non-fictional CEO of Facebook.

But in virtue of what is the self-considered-as-phenomenal the same thing as the self-considered-as-noumenal? With *The Social Network*, there are several factors that make it clear that the Mark Zuckerberg of the film is the same person as the CEO of Facebook: they have the same name, there is significant resemblance between the properties and history of the two, and the film-makers clearly intended the film to be about the CEO of Facebook. But nothing along these lines would help connect the phenomenal self with the noumenal self: the noumenal self does not have a name, has radically different properties from the phenomenal self, and there is nobody who intends the phenomenal self to be the same as the noumenal self.³⁹

None of this shows that that there is no relation between the phenomenal and noumenal selves that might unite them by making the phenomenal *about* the latter. After all, there are textual reasons to say that noumenal self *grounds* the phenomenal self, affects the latter, etc. But, again, these relations are ones that hold quite generally between phenomena and noumena,⁴⁰ so they if are enough to unify the self, they will unify phenomena and noumena more generally.

³⁹ These same points apply to Wilfred Sellars' suggestion that the phenomenal/noumenal distinction is the Cartesian objective/formal distinction (Sellars, Wilfred: *Science and Metaphysics: Variations on Kantian Themes*. London. 1968, 31). For Descartes himself sets limits to how far the objective reality and formal reality of an object can diverge (cf. the discussion of the most universal things in Meditation I, and the argument for God's existence in Meditation III).

⁴⁰ This is especially clear in the Fourth Paralogism in the A edition (e.g. KrV A379-80), even though this is a passage often cited in support of the two-world reading.

Nothing requires a two-world interpreter to follow Van Cleve in making appearances and things-in-themselves so radically different. I see no principled way to show that there could not be a two-world reading that can answer the challenge I have described. But it is clear that the basic premise of the two-world reading puts it in a difficult position. The strength of Van Cleve's virtual-objects approach is that it shows how appearances and things-in-themselves cannot be the same things – indeed, this is part of what makes his two-world reading admirably clear. Yet this same feature of this interpretation forces him to reject the unity of the self. If a two-world interpreter departs from Van Cleve, however, and attempts to unify the self with a weaker understanding of the distinction or with any of the obvious relations between the phenomenal and noumenal selves to unify them (substance/attribute, cause/effect, ground/grounded), then the view will collapse into a one-world view.⁴¹

That completes my primary defense of the One-Self/One-World Argument. There are, I think, two initially plausible replies a two-world interpreter might make. These are the topic of the next section.

Two Replies

Reply 1: Unique features of the self

A two-world interpreter might respond to the above by claiming that there is something unique about the self that makes for a unified phenomenal-noumenal entity in that case alone. In general, this line of response must contend with the second and third premises of the One-Self/One-World Argument: since Kant repeatedly treats the appearance/in-itself distinction as applying univocally and uniformly, this suggests there is nothing distinctive about how the distinction applies to selves. In addition, Kant begins all of his major works with the assumption of a single self firmly in place. If the two-world view were correct, but Kant had some reason for making the self an exception (in being the only case of a thing that is both phenomenal and noumenal), we would expect him to explicitly introduce and motivate that exception somewhere. But he does not.

Since a two-world interpreter might insist that the interpretation has enough virtues to make up for those textual problems, however, we should consider how far this reply might go

⁴¹ Though my focus here has been metaphysical, there are difficult epistemological issues in the vicinity. Given Kant's epistemological strictures, we should ask how he would be *entitled to make claims* about the connections between noumena and phenomena (cf. Klemme, *Kants Philosophie des Subjekts*, 245-270).

if we set them aside. To be clear, what is needed is some basis for saying that the phenomenal self is united to the noumenal self that does not imply that phenomenal things generally are united to noumenal things. Moreover, it would seem that (unlike Schopenhauer) Kant thinks that different phenomenal selves are united to different noumenal selves, and the reply needs to make room for that. It would not do to, say, to claim that a given phenomenal self is united to a certain noumenal self because they both have some structural feature, unless that structural feature is had *only* by that particular phenomenal self and that particular noumenal self. Otherwise, it could immediately follow that have all phenomenal selves have the same noumenal self.

There are no absolute philosophical obstacles to making this reply work within those constraints. The most straightforward approach would be for a two-world interpreter to say that, for each phenomenal self, it is a brute fact that it is united to certain noumenal self in a way that phenomenal things other than selves are never united to noumena. A two-world interpretation that appealed to such brute facts would be internally coherent. It would, however, be mysterious and ad hoc.⁴²

It would be more satisfying if the two-world interpreter could identify features of the phenomenal and noumenal selves *in virtue of which* they were united. For the above reasons, it will not due to appeal to the mere fact that the phenomenal and noumenal selves have some causal, inherence, or grounding relation, so the reply would need to pick out a more specific relation that holds only between particular phenomenal and noumenal selves.

Since rational selves are very distinctive entities in Kant's philosophy, there are many specific relations that could potentially be appealed to. To my mind, the most promising approach would be to appeal to self-affection, which seems to be (within Kant's theoretical philosophy) the distinguishing feature of selves that most directly concerns the phenomenal/noumenal distinction.⁴³ Put in the terms of one version of the two-world reading,

⁴² One might be drawn to this line if one thought, perhaps in a Hegelian spirit, that Kant's theory assumed but failed to explain the unity of the self (cf. Düsing, Klaus: "Constitution and Structure of Self-Identity: Kant's Theory of Apperception and Hegel's Criticism," *Midwest Studies in Philosophy* 8, 1983, 409-31).

⁴³ See, for instance, §24 of the B Deduction. This approach was suggested to me by Béatrice Longuenesse. Other possibility would be to appeal to the unity of apperception (Jauernig, *How to Think about Things in Themselves* considers this). I find it harder to see how that approach would work, though, and since Kant's introduces the unity of apperception in the *KrV* long after he introduces self-affection and his idealism, it is even less likely that the unity of apperception is supposed to play unique a role in unifying the phenomenal and noumenal selves. Finally, a two-world reader might think to appeal to the transcendental I (cf. *KrV* B157-159)

here is a principle would have the relevant implication: a particular phenomenon P is united to a particular noumenon N if and only if (a) P exists in virtue of intuitions that arise from N's affection and (b) those intuitions inhere in N. Because this principle does not explicitly appeal to selves, it is not obviously ad hoc. Conditions (a) and (b) hold only for particular phenomenal selves and noumenal selves, since with other empirical intuitions, the noumenon responsible for the intuitions is presumably distinct from the noumenon in which the intuition inheres. For that reason, the principle implies that the self will be the only case where a phenomenon is united to a noumenon.

Even this form of the reply, however, would implicitly appeal to a brute fact: that (a) and (b) are necessary and sufficient for unification. (a) and (b) describe an intimate metaphysical relation, but one that sounds like less than sameness. As an analogy: a tattoo artist gives herself a tattoo of an animal. The tattoo was produced by her and inheres in her, but there is no intuitive temptation to say that the tattoo is the same thing as her (unless we think inherence is generally sufficient for sameness). This shows that the above principle relies on a counter-intuitive brute fact. I suspect that the same will be true of any attempt to unite the phenomenal and noumenal selves on the basis of some specific relation: there will be some implicit but counter-intuitive brute insistence that that relation is necessary and sufficient for sameness. Without such an insistence, though, this reply has nothing to say to the philosophical motivation behind the third premise of the One Self/One World Argument (the metaphysical uniformity of the general distinction with respect to sameness).

Moreover, the above-mentioned textual problems this line faces are at least as serious. Not only does Kant frequently write as though the phenomenal/noumenal distinction were uniform, but he seems to assume a unified self before he introduces any distinctive features of the self that might explain that unity.

Reply 2: Not enough noumena

A second potential response to the One-Self/One-World Argument might be based on Paul Guyer's suggestion that the only noumena Kant postulates are God and the soul.⁴⁴ If the soul and God were the only entities with noumenal aspects or components, then (setting God

to unify the self, but if so, she would need to explain both why that would not unite all rational selves into one and why the transcendental object (cf. KrV A109) does not unify other phenomena and noumena.

⁴⁴ Guyer, Paul: *Kant and the Claims of Knowledge*. New York. 1987, 334-5. Context suggests Guyer must have in mind the specifically noumenal soul. Guyer now explicitly endorses a two world view (Guyer, Paul: *Kant*. New York. 2006, 69-70).

aside) one could say that while the empirical self is united with the noumenal soul, there are no corresponding noumenal entities with which other empirical entities could be united. Such a claim would then seem to allow one to appeal to general relations in unifying the self (e.g. substance/accident) without having to worry about a general one-world interpretation resulting.

The principal difficulty this reply faces is textual. Kant explicitly claims that there are things-in-themselves for all appearances (e.g. KrV Bxxvi, A252, B306).⁴⁵ Even if the relevant passages can be dealt with, Kant would need to be read as positively denying that there are any things-in-themselves, as opposed to remaining agnostic. Otherwise, the resulting view would be a sort of agnostic one-world view.

There would be additional difficulties for this reply in avoiding the unification of *all* phenomena with either the soul or with God. For phenomena generally may stand in many of the same relations to God and the soul as the phenomenal self stands to the noumenal self, but having all phenomena as closely united to the noumenal self as the phenomenal self is, or having them all form one thing with God (à la Spinoza) would surely be an un-Kantian result.

Conclusion

The One-Self/One World Argument supports a general interpretive principle: since the self is an instance of a uniform appearance/thing-in-itself distinction, the relationship between the phenomenal and noumenal selves should be like that of phenomena and noumena generally unless there is a principled reason to think otherwise. While I have applied this principle to the two-world interpretation, I think the principle also yields constraints on how a one-world interpreter should understand ‘same thing’ relation. In particular, I think the noumenal and phenomenal selves cannot be identified, since they can have distinct determinate causal features,⁴⁶ and the above principle therefore suggests that no phenomena are identical to noumena. A one-world interpreter must therefore, I hold, spell out the ‘same thing’ relation as something other than identity.

⁴⁵ For a useful discussion, see Allais, Lucy: “Transcendental Idealism and Metaphysics: Kant’s Commitment to Things in Themselves”. In *Kant Yearbook: Metaphysics*. 2010, 1-32.

⁴⁶ Cf. KrV A532/B560ff. This is a point sometimes made by two-world interpreters (e.g. Guyer, *Kant*, 68), though typically without realizing that the general principle applies to the self. This issue is perhaps part of why Allison avoids putting his interpretation in terms of identity (cf. Allison, Henry: *Idealism and Freedom*. Cambridge. 1996, 11-12).

How should two-world interpreters respond to the One-Self/One-World Argument? I believe that the best response is neither to challenge the validity of the argument, nor to simply charge Kant with inconsistency. Instead, two-world interpreters should reject the first premise. To do so, she must find some other way of reading the passages where Kant seems to suggest there is one self with noumenal and phenomenal sides, and find ways of making sense of those aspects of Kant's system that make appeal to the single self. This is no small task, since there are many such passages, and the unified self plays a key role throughout Kant's philosophy.⁴⁷

⁴⁷ Thanks to Ralf Bader, Andrew Chignell, Dana Goswick, Stefanie Grüne, Anja Jauernig, Béatrice Longuenesse, Mike Raven, Sagar Sanyal, Jonathan Simon, and Nick Stang. Several years ago, I ran a crude version of this argument by someone at Dojo Restaurant in New York. That person encouraged me to develop the idea, which it has taken me a while to do. I'm very grateful to him or her, but cannot remember who it was!