Is Kant's Concept *World* Self-Contradictory?

Aaron Wells, Metropolitan State University of Denver

Final version forthcoming in *Archiv für Geschichte der Philosophie*
Please cite published version when available

Abstract: There is a relatively overlooked problem with Kant's claim that in inquiry we must treat nature as if it were actually infinite in extent: he also states that the concept of an actually infinite world is self-contradictory. This threatens to make the command to treat the world as infinite incoherent; the problem also affects his error theory for traditional metaphysics and his account of the sublime. After laying out this worry in greater detail than has been done so far, I consider and reject readings on which Kant either does not consider this case a violation of ought-implies-can or allows that in this case, following an incoherent demand is epistemically acceptable on consequentialist grounds. Then I turn to my preferred solution, which takes the relevant regulative principle to concern Kant's noumena (understood as intensionally distinct from things in themselves), whereas the contradiction is provoked only by representations of phenomena.

Kant's assertion that specific and substantive norms for inquiry flow from the interests of reason is influential and controversial. Among other things, he holds that scientific inquiry broadly construed ought to consider its objects *as if* they had certain properties—nature is to be treated as if it were infinite, the soul is to be treated as if it were a simple substance, etc.—without positively asserting *that* they have these properties. Equally challenging is his suggestion that the illusions of traditional metaphysics spring from the very same interests of reason. Despite extensive discussion,[1] relatively little attention has been paid to a looming inconsistency between Kant's view that some concepts of traditional metaphysics are self-contradictory, and his apparent employment of those very concepts in norms of inquiry.

In brief, Kant finds contradiction in the concept of a world that is completely given in its spatial and temporal properties. Yet the concept of an infinite world also appears in one of Kant's regulative rules for inquiry. To grasp this rule determinately, Kant suggests that we must conceive of an object that, if it existed, would fully satisfy the rule. Troublingly, this seems to require conceiving of a world that is infinite in spatial and temporal extent. Now, if a necessary means for a regulative rule to state a goal for inquiry is through the representation of an object, then that object had better be logically consistent. But the concept of a completely given spatiotemporal world *is* self-contradictory, making it unable to guide inquiry. Nor can the problem be confined to the case of regulative rules. As we will see, it threatens to infect Kant's explanation of the errors of dogmatic metaphysics, and even his account of the moral significance of the sublime.

In this article I first lay out these tensions in detail (Section 1). Next I consider, but ultimately reject, three proposals that target the normative status of Kant's regulative principles (Section 2). The first of these focuses on Kant's restriction of his ought-implies-can principle to what concerns the will as such, as opposed to circumstances and inabilities that are outside. But this does not solve the problem with self-contradictory principles: if these

[1] Recent entry points into the large literature on these topics include Kraus (2025), Proops (2021), Schafer (2023), and Willaschek (2018).

principles are followed voluntarily, then one's inability to follow them stems from their self-contradictory character, which is internal to the will. A second proposal, inspired by a suggestion from Robert Stern, would simply weaken the ought-implies-can principle to extend beyond what ordinary humans can do. While I think such a move is promising elsewhere, it doesn't help with the threat of contradiction. The third proposal would, against the grain of many recent readings, take Kant as endorsing a localized type of epistemic consequentialism, specifically regarding regulative principles of pure reason. The idea would be that even if an object of inquiry is self-contradictory, it might be in our best interest to let this contradiction remain hidden and follow the relevant regulative rule as if it were consistent. Although the importance of epistemic consequences to Kant's account of regulative principles deserves further scrutiny, I'll argue that given his views on unfulfillable desires and the nature of rational imperatives, an appeal to epistemic consequentialism won't resolve the relevant incoherence.

I then defend a response on which the self-contradictory concept of the world is not identical to the ideas of reason that play a positive role in Kant's system (Section 3). I propose that the self-contradictory concept is of the world as a phenomenon, and the regulative concept is of the world as a noumenon. Only the latter concept can consistently attribute absolute properties to a world. Among other possible objections, I consider how a representation of the world as noumenon could regulate inquiry into the phenomenal world, given that concepts of noumena precisely lack spatiotemporal content. The noumenal representation posits a global property of a world—its consisting of an actual infinity of objects—that is not inherently spatiotemporal, and *not* realizable in any phenomenal world. It just for this reason that the representation of the world as noumenon is both logically consistent and plays a regulative role for inquiry that no empirical concept could substitute for.

1. A self-contradictory concept of the world

To set the stage, I'll give a preliminary explanation of Kant's claim that a certain concept of the world is self-contradictory. I'll then lay out three apparent tensions between this claim and his other commitments.

Kant's claim about self-contradiction appears in his Antinomies, which set up paired indirect arguments for opposed metaphysical theses about the spatiotemporal world. Two so-called mathematical Antinomies are disputes about the world's temporal and spatial extent and the divisibility of its material parts. To keep things simple, I focus on the size of the world, not its divisibility. For this Antinomy, the opposed theses do not turn out as contradictories in the traditional Aristotelian sense. That would require that one proposition be true and the other false, but by Kant's lights, the opposed claims of the mathematical Antinomies are *both* false.[2]

These propositions are both false because they "have as their ground an impossible concept of the object."[3] An impossible concept is just a self-contradictory concept: these propositions attempt to predicate a self-contradictory concept of an object. The concept that causes trouble is, roughly, the concept of the world as endowed with absolute spatial and temporal properties.[4] To be more precise, the specific concept of world involved here need not

[2] See 20:291.

[3] A793/B821; also see A431/B469; A740/B768; 18:404.

[4] Kant vacillates between talking about a single contradictory concept of the world and a cluster of "world-concepts" that are all contradictory for the same reason (A420/B447). I use the singular here because I focus on the first Antinomy, which deals with one such world-concept. Another point of clarification: the self-contradictory concept need not be that of the "world…given in its totality" (A793/B821). Absolute spatiotemporal properties can also be attributed to spatial or temporal *parts* of the world. Given a finite material thing, one can ask whether it is infinitely divisible, with no need to consider the world as totality (A415/B443;

refer to all things in general, and abstracts from qualitative and causal properties. The world so construed is "the mathematical whole of all appearances," insofar as these material appearances are aggregated "in space or time so as to bring about a quantity," which would be an "absolute totality."[5] 'Mathematical' is a term of art: Kant is not referring to the infinite space underwriting Euclidean geometry, which he describes as subjective, but to the aggregate of extended objects making up the world.[6] The concept of the objective world as a totality, then, in some sense involves a contradiction.

When it is asked whether the time that has elapsed in the world so far is infinite or finite, at first glance each possible answer looks internally consistent. But Kant insists that *both* options are self-contradictory. Dogmatic metaphysicians only advance the conflicting options because they are in each case committed to an underlying self-contradictory concept as the "condition" for the judgments they endorse.[7] To attribute a "fully elapsed time" to the world, whether infinite *or* finite, is already to attribute an absolute property to it, and thereby to try to predicate a self-contradictory concept of the world.[8]

For now, I set aside further discussion of why this concept is inconsistent (on which see Section 3.2). I instead consider what it means for the concept to be *presupposed* in the First Antinomy. Both thesis and antithesis assume this concept, which I'll call *W(a)*. The concept is used to construct paired reductio arguments for the conclusions that the world is finite and that it is infinite. Kant maintains that these arguments are valid, and I'll grant this to avoid going into unnecessary detail. If there is no problem with the arguments themselves, the roots of contradiction must lie in the underlying concept *W(a)*, as he explains with the following example:

> The two propositions: a square circle is round, and: a square circle is not round, are both false. For, as regards the first, it is false that the aforementioned circle is round, since it is square; but it is also false that it is not round, i.e., has corners, since it is a circle. The logical mark of the impossibility of a concept consists, then, in this: that under the presupposition of this concept, two contradictory propositions would be false simultaneously.[9]

The truth-value of a proposition, this passage suggests, partly depends on its subject-concept. If a self-contradictory concept is used as the subject of a proposition, then both positive and negative predications of that subject term will turn out false. And Kant holds that the concept *W(a)* is self-contradictory, just like the concept of a square circle. Thus there is an internal contradiction in the proposition that the world is given in its total temporal extent as infinite. The contradiction arises because the proposition needs *W(a)*, the concept of the world as given in its total extent, as its subject term in order to predicate finitude or infinitude of it. But this concept is self-contradictory.

This conclusion, as Ian Proops and Anja Jauernig have noted, threatens to undermine some of Kant's other commitments. I'll expand on their suggestions, detailing three problems that threaten if this concept is self-contradictory.

A526/B554). As Proops (2021, 275) points out, the same goes for the world's temporal extent: to say the world is eternal *before some specified time* is to commit to completed temporal infinity.

[5] A419–20/B446–47; A413/B440.

[6] The subjective space presupposed by geometry is "not a real object" (A429/B457n.; see also B40, A431 /B459, A715/B743; 4:506–7; 20:420; 28:568; 29:837; 29:864; Winegar 2022).

[7] A503/B531; also see A505–6/B533–34 and 18:401–403. This tells against Allison's (2004, 443–44) suggestion that contradiction can be avoided by picking either the finite or infinite option.

[8] A410/B437.

[9] 4:341.

First, Kant apparently needs *W(a)* to define a principle he takes to govern empirical inquiry, so there will be trouble if the concept is self-contradictory. Here is the relevant principle:

> We have to pursue the conditions of the inner as well as the outer appearances of nature through an investigation that will nowhere be completed, *as if* nature were infinite in itself and without a first or supreme member.[10]

For the question of the spatial size of the world, the "conditions" in question are larger spaces as conditions for bounding smaller spaces they contain.[11] The principle says we should inquire as if there exist no "first" or simple parts, and as if there is no "supreme" space that contains all other spaces. Some concept of a world as an "absolute totality," infinite in certain respects, is by Kant's lights needed for us to posit this "goal" for inquiry.[12] A consistent concept of an object is needed because

> in order to represent [regulative] principles determinately, reason conceives of them as the cognition of an object, cognition of which is completely determined with respect to these rules—though the object is only an idea—so as to bring cognition through the understanding as close as possible to the completeness that this idea signifies.[13]

Only with the intentional object of this concept can we determinately think the actual spatiotemporal world *as if* it were given in its total extent as infinite. The problem is that the concept in question looks to be an instance of the concept *W(a)*, but *W(a)* is self-contradictory. A regulative principle commands that we do something. So if it relies on *W(a)*, then it is analogous to a command to act as if the intentional object of the concept of a square circle existed.[14] If ought implies can, then since the principle commands nothing and cannot be followed, we also cannot be obligated to follow it. In brief: a self-contradictory principle can have no regulative force.

A related second problem arises for Kant's diagnosis of what he calls transcendental illusion. He wants to give an error theory for the claims of dogmatic metaphysics. He thinks we are apt to form the false belief that theoretical reason's ideas, or their intentional objects, can be instantiated in experience. Dogmatic metaphysicians then think they can use these ideas to grasp the world in itself. This illusion is not merely contingent or pathological, but relates to constitutive features of theoretical reason. There is some disagreement about how this works, but Proops's reconstruction should suffice as an illustration. On this reading, an essential part of transcendental illusion lies in conflating the following two principles:

> (*Necessary Injunction*) "You must seek, for the object of any cognition, the series constituting its unconditioned (real) condition."
>
> (*Necessary Existence*) "Necessarily, for the object of any cognition, the series constituting its unconditioned (real) condition exists."[15]

---

[10] A672/B700. Also see A508–9/B536–37; A685/B713.

[11] Respectively, A685/B713 and A412/B438–9. I focus on space here because it is more straightforward than time. Kant does not think a space could have a well-defined boundary without being contained in a larger space, so the containing space conditions the space it contains.

[12] A685/B713; A644/B672. Though an object instantiating the concept in question "cannot be given in any experience," it is "produced…in the idea," or represented as an intentional object (A510/B538).

[13] 4:331–332.

[14] For similar points see Proops (2021, 425; 461) and Jauernig (2021, 334)

[15] Proops (2021, 49); also see Hogan (2021, 433–39). Stang (2016, 280–96), Schafer (2023, 159–68) and others think Kant endorses something like *Necessary Existence*. I pass over this debate, since my focus is on *Necessary Injunction*.

We take *Necessary Injunction* to be a rational demand, and are warranted in doing so. *Necessary Injunction* is part of reason's natural and proper functioning, and so "must contribute to the perfection" of inquiry, rather than merely "confuse" it.[16] However, we tend to slip into the unwarranted further assumption that the goal it demands for inquiry actually exists—in other words, we also tend to also endorse *Necessary Existence*. In the case of the mathematical Antinomies, we conflate the regulative demand to seek larger and larger spaces and times (or smaller and smaller parts), and the factual conclusion that infinite space and time (or simple monads) exist. This conflation is illusory, but it starts with the rational plausibility of *Necessary Injunction*. Rationally plausible principles must at least be logically consistent. So it's a problem for Kant's error theory if the injunction in question prescribes a self-contradictory goal. One might be tempted to respond that dogmatic metaphysicians overlook these subtle contradictions. Even if that's true, readers of the *Critique* undergo what Proops calls a "loss of innocence," unveiling the contradiction in *W(a)*.[17] This is a problem because Kant's error theory dictates that transcendental illusion, like perceptual illusion, is cognitively impenetrable. Becoming aware of the illusion cannot make it go away, since a propensity to it is bound up with reason's proper functioning.[18] Error results not from reason itself, but from mistaken judgments, for example confusion of *Necessary Injunction* with *Necessary Existence*. If the former is revealed to invoke a self-contradictory concept, its vacuity will become obvious. Then the illusion is not cognitively impenetrable after all, undermining Kant's error theory.

The third problem lies further afield, in the aesthetic theory of the third *Critique*, and specifically in Kant's account of how an idea of the world as a completed magnitude evokes what he calls the *mathematical sublime*. The feeling of the sublime comes from a conflict between two cognitive capacities. On the one hand, the imagination strives to synthesize infinite spaces and times step by step. But given our temporal finitude, this can never be achieved. On the other hand, pure reason presses us to transcend these limits. "The voice of reason…demands" that "even the infinite" be presented as a completed totality.[19] This demand applies not just to space and time, which we do not perceive directly, but also to material things *in* space and time.[20] The sublime's unfulfillable demand has aesthetic and moral weight: it "demonstrates" that we possess a "supersensible faculty" surpassing our spatiotemporal finitude.[21] The problem raised for the sublime by the self-contradiction in the idea of the world is that reason's demand for absolute totality must be logically coherent, even if it is not a demand we can meet in practice. For only if this demand is coherent can the sublime have affective force and teach us truths about our own rational nature. A demand for a self-contradictory goal could not express anything, let alone our "supersensible vocation."[22]

I conclude this section with a point of clarification. I have focused on the special case of apparent contradictions in alleged rational principles. So I am not considering the *general* question whether Kant allows any thinkable contradictions, and will remain neutral on that. For context, though, I will sketch a recent debate on the topic. Kant sometimes writes that

---

[16] 4:331; also see A642–3/B670–1.

[17] Proops (2021, 460).

[18] On rational illusion and proper functioning, see A642–43/B670–61. For a comparison with perceptual illusion, where appearances remain the same even if we are "not deceived," see A297/B354.

[19] 5:254. Schafer (2023, 130–37) helpfully situates the mathematical sublime within Kant's broader account of reason and comprehension. Also see Winegar (2022, 654–57).

[20] Reason's demand is applied to "appearances" at 5:255.

[21] 5:250.

[22] 5:257; 5:267.

self-contradictory thoughts are "nothing."[23] This might seem to entail that there are no self-contradictory thoughts, but not all interpreters are convinced. For one thing, he may only hold that contradictions only count as nothing for *pure general logic*, a subdiscipline that states necessary, non-empirical laws of thinking. Contradictory statements, as aberrations from the laws, are automatically excluded. Logic's impure side can study the sources of error—including psychological failure to detect implicit contradictions.[24] So one strategy for accommodating the representation of contradictions would be to say that they can erroneously be represented *as* thoughts, though they are in fact pseudo-thoughts, outside the scope of pure general logic. This approach has the shortcoming of explaining only pseudo-thoughts that fuel error. So it does not cover an important case raised by Kant's acceptance of reductio arguments in mathematics and natural science. Contradictory premises in these arguments are genuine thoughts, not pseudo-thoughts, since they have valid inferential consequences.[25] Kant also states directly that in some sense we can think contradictions, such as that a triangle has six sides.[26] To accommodate these cases, one might stress his suggestions that the problem with contradictory thoughts is that they intrinsically fail to represent a possible object, so are automatically false. This might be all he means in saying that contradictory propositions are nothing. The principle of contradiction could be seen as a rule governing the truth of logical contents, rather than a structural rule of those contents.[27]

My aim is not to adjudicate among these readings, but just to point out that neither way of accommodating contradictions can help with the *specific* problem under discussion. What's at issue are not mere pseudo-thoughts or formally false propositions, but allegedly authoritative rational principles. Even if these principles are, in some bare sense, thinkable, the question at issue is whether they can coherently command us to do anything. And that is something contradictory principles cannot do.

2. Ought-Implies-Can and Consequences

A first line of response to the threat of incoherence would focus not on the idea of the world itself, but on the scope and implications of Kant's ought-implies-can (OIC) principle. I consider three variations on this response. The first two would directly weaken the OIC principle. The third reads Kant as, in limited domains, an epistemic consequentialist. I think these issues deserve further attention, but I conclude that they do not afford a good response to the current problem.

Any restriction on OIC might seem like an uphill battle. Kant famously accepts such a principle for *unconditional* oughts, or what we would now call all-things-considered moral obligations.[28] Proops argues that Kant also takes an OIC principle to govern epistemic oughts imposed by the regulative rules of theoretical reason. There is some textual support for this. When Kant writes that "reason cannot command the pursuit of an end which is known to be

---

[23] Bxxviii; also A150/B190; A254–55/B310; 4:341; 8:195; 29:792.

[24] A54/B79. On hidden or implicit contradictions see 28:544, 29:965, and also Schafer (2023, 103–104).

[25] As Leech (2017) and Proops (2021, 460) observe. See e.g. A792/B820; 9:52; 24:233; 24:893.

[26] 24:719. This judgment is not asserted as true, only *posited*, so it does not count as a proposition (*Satz*) in Kant's sense. For a different example, compare 8:194(n.).

[27] See Leech (2017, 358–59) and Lu-Adler (2017) for proposals along these lines. Structural rules for propositions would be analogous to the syntactic and semantic rules that must be followed for sentences to be more than mere nonsense. There are many passages where Kant describes contradictory thoughts as false (A59–60/B84; 9:51; 24:826; 28:544; 29:960–61). To evaluate a thought as false rather than true, one presumably must be able to think it.

[28] See for example 6:380: a human being "must judge that he can do what the [moral] law tells him unconditionally that he ought to do."

nothing but a phantom of the mind," for example, he makes no explicit restriction to unconditioned moral ends.[29]

This extension of OIC to theoretical contexts also has common-sense appeal. For example, suppose my reasoning invokes a mathematical assumption that is necessarily false, even though I believe it to be true. If I'm asked to prove this assumption—for example, the general comprehension principle in naïve set theory—then I am asked to do something impossible. I might be conditionally obligated to provide a proof if there is one, or have a disjunctive obligation to either provide a proof or give up the assumption. But how could I be under an unqualified theoretical obligation to give such a proof?

Still, one might resist attributing a theoretical OIC principle to Kant. A first reason is that he sometimes indicates that OIC "only" concerns "the will," apart from whether the agent can cause further desired effects.[30] The regulative principles of theoretical reason do not only concern the will. At minimum, they also require successful representation *as if* things were such and such a way. So theoretical regulative principles may not fall under OIC, because some contingent inability to form the appropriate representations might prevent an agent from following them. This does not obviously help with the problem of self-contradictory regulative principles, however. One cannot consciously will to follow such principles, just because they are contradictory. This inability is internal to the will, not due to contingent circumstance. So it is no exception to OIC. A complication is that since Kant allows hidden contradictions, there could in principle be a gap between what an agent takes herself to be (consistently) willing, and what she is really committed to (inconsistently) willing. But since the epistemic consequentialist line of thought discussed below will try to exploit this gap, I leave it aside for now.

One could instead target a common line of support for OIC principles, namely that it's incoherent to take an action to be obligatory if no agent is even capable of performing it, because an agent cannot be blamed for failing to do something she cannot do. Against this, Robert Stern has argued that a strong OIC principle is not needed to ensure obligation. For Stern, one needs only the weaker assumption that an agent sufficiently like us—though perhaps superhuman in some ways—could do as the obligation prescribes.[31]

Even if Kant would not accept this weakening in the moral domain, he might do so for other obligations. In the case of the sublime, we saw that reason demands a goal that our other faculties constitutively cannot achieve. This demand remains in force even if our finitude prevents us from fully carrying it out. He may even think that the abilities of a conceivable superhuman agent are relevant to our epistemic obligations, as Stern suggests. For Kant, we must pursue the goal of completely explaining the material world mechanically, despite his admonition that we will never fully explain even a single blade of grass in this way. Yet a superhuman understanding "might be able to" accomplish this sort of explanation, and he seems to think this possibility somehow underwrites *our* epistemic obligation to pursue mechanical explanations.[32]

---

[29] 5:472; cf. A314/B370–71. The broader context of the passage, especially 5:471(n.), could however be taken to show that he has moral ends in mind.

[30] 5:20–21. One of Kant's examples is the moral obligation not to make a lying promise. If the world is uncooperative, it may turn out for all practical purposes, an agent *cannot* fulfil some promise. Since OIC concerns the will, however, these external circumstances do not release the agent from moral obligation.

[31] Stern (2004, 48). Kant notably grants the possibility of ethical *assistance* by some "higher moral being," without taking this to undermine our moral obligations (6:98).

[32] 5:405–406. One way of reading the reference to a superhuman understanding is helping establish that it is possible for us to keep pursuing mechanistic explanations indefinitely. If this were not possible, we could not be properly "required" to commit to it (5:388).

Even if this is granted, Stern's proposal would not resolve the problem of a self-contradictory regulative idea. Here, Kant apparently takes there to be an epistemic obligation to seek a self-contradictory object as a goal. The problem is not one of human foibles or finitude: *no* conceivable agent could carry out this goal.

So I turn to a different line of thought about why a specifically epistemic OIC principle might fail. This is a reading of Kant as, in some contexts, an epistemic consequentialist. While I only entertain rather than seek to adopt this reading, I will argue that it is surprisingly credible, raising issues that deserve further attention in the literature. However, I aim to show that it is not relevant to the OIC principle in question. If Kant is an epistemic consequentialist, this could only hold in some restricted areas, which will not include the regulative principles at issue. No help with our initial worries can be found here.

A brief introduction to epistemic consequentialism first. Many philosophers now agree that we have epistemic ends, despite disagreement on which ends these are. This raises the possibility of epistemic trade-offs. For example, if our epistemic ends are to acquire as many true beliefs as possible, but a necessary means to acquiring the greatest number of true beliefs is to first adopt a solitary false belief, then it is rational to cause ourselves to hold this one false belief. The false belief in question might turn out to be an overly optimistic belief about our abilities. Then why couldn't this false belief be the belief that we are able to investigate nature as if it were infinite in itself in spatial and temporal extent? If so, we epistemically ought to have this false belief about our abilities, and ought to investigate nature as if it were infinite in itself. Kant allows that contradictions can be hidden, and the proposition that the world is eternal is not obviously contradictory. So there seems to be no in-principle barrier to allowing an epistemically advantageous belief to be not just false, but self-contradictory. From an internal perspective, it will seem to us as if there is no ought-implies-can violation, since we believe we can follow this epistemic norm. So long as the contradiction remains hidden, we will only draw a limited set of consequences from it (even if as contradictory, it in fact entails anything whatsoever). But there really is an ought-implies-can violation, since we are unable to follow the norm. This violation might nonetheless be acceptable, even rational, on epistemic consequentialist grounds. In this case, apparent irrationality might turn out on balance rational.

But wouldn't Kant reject this rather crude epistemic consequentialism?[33] Certainly, he could not accept it across the board. He seems willing, however, to countenance false beliefs about our own powers or abilities, so long as these false beliefs serve some end:

> If we were not determined to apply our powers by the representation of an object until we had made sure of the adequacy of our capacity for producing it, then the latter would remain mostly unused. For we commonly learn to know our powers only by trying them out. Nature has therefore combined the determination of our power with the representation of the object even prior to knowledge of our capacity.[34]

In order to develop and use our abilities, it is advantageous for us to form beliefs that we are capable of carrying out our ends, even when the evidence for this is objectively insufficient. Some of these beliefs will turn out false. So our desires for the ends the beliefs are about will be idle. Yet these idle desires "incite us…to activity" and further our overall epistemic ends.[35] While it's rational for us to moderate this belief-forming tendency, Kant thinks it would be

---

[33] As Schafer (2023, 9–10) summarizes, Kant is widely taken as an epistemic non-consequentialist, whose "requirements of theoretical reason can be derived from...some form of the categorical imperative." By contrast, Vaihinger (1924, 30) thought Kant was committed to accepting even self-contradictory principles if they have sufficiently good consequences.

[34] 20:231; also see 5:177–78. For more on these passages, see Englert (2017).

[35] 29:895.

irrational to try to eliminate it, because the activity it incites is good. We thereby have warranted beliefs about our own abilities that nonetheless often turn out false.[36]

This belief-forming tendency might seem limited to our animal nature. It depends on empirical facts about how we develop our epistemic powers, rather than flowing from the nature of reason itself. However, one could argue that since for Kant the epistemic goods of theoretical reason are means to unconditioned moral ends, there is also a means-ends structure among the epistemic goods of theoretical rationality, such that holding false beliefs might serve as a means.

For a helpful contrast, consider his famous comparison between the unconditioned good and a jewel that shines all by itself. An unconditioned good has intrinsic value even if it brings about no good consequences whatsoever. But the regulative principles of theoretical reason are not like this. They derive at least some of their normative force from their "use" or "goal," which is systematic unity for our theoretical inquiries.[37] If we were entirely unable to bring about this goal by way of regulative principles, then these principles would have no positive value, or at least much less positive value. Their value is, in whole or in part, conditional. Kant explains that the reason why "I must accept" regulative principles is a matter of their *subjective* rather than objective sufficiency: if I don't accept them, "I make no progress" in inquiry.[38] The conditional character of regulative principles also flows from his view that in purely theoretical matters, there cannot be objective duties—from which he concludes that there are no duties to have theoretical commitments as such.[39] It follows that there are no objective duties to represent *as-if* via ideas of reason, unless there are practical grounds for doing so. This need not exclude some additional, intrinsic goodness for regulative principles, that might ground their normativity. The key point is that valuable consequences for inquiry confer at least some of these principles' normative force.

For further evidence, consider Kant's surprising argument that even dogmas of traditional metaphysics might advance the epistemic end of "pure rational unity," understood as an interest of theoretical reason (A475/B503). These consequences give us defeasible warrant for adopting rationalist dogmas over their disunified rivals. We then have prima facie grounds to adopt a positive epistemic attitude—such as doctrinal *Glaube*—toward these dogmatic theses. Several readings are then possible here. A weaker reading would hasten to add that in the event, these grounds are defeated because speculative metaphysics is not the only means to the end of pure rational unity. Regulative principles, and better still pure practical reason, are shown to give alternative means to the same end of pure rational unity. Moreover, in the case of the spatial and temporal extent of the world, the proposals of dogmatic metaphysics are shown to not even be logically consistent, which overrides any positive interest of reason in unity. Dogmatic metaphysics is therefore dispensable. On a stronger reading, these epistemic grounds are not always defeated: Kant endorses not just

---

[36] I use 'warrant' as neutral between (internalist) justification and (externalist) entitlement. I do so because as we'll see, Kant contrasts the human case here with non-human animals, and warrant can cover both cases. I use 'belief' roughly as it's used in contemporary epistemology—leaving aside the debate over whether epistemic consequentialist considerations should only lead to weaker attitudes such as acceptance. I leave untranslated Kant's technical term '*Glaube*.'

[37] A644/B672. Kant elsewhere writes that these principles serve only to "point the understanding in the right direction" (A323/B380). Note that I am not ruling out unconditional epistemic obligations in Kant, as defended for example by Hadisi (2022). I also do not exclude unconditional goods of theoretical reason *tout court*, though I think there are grounds for skepticism about this (see e.g. 4:444; 5:142(n.); 5:442; 27:1322). I'm just saying that *regulative principles* of theoretical reason in particular are good because of their consequences. So they can't impose unconditional obligations, just because they are means to reason's ends (see also Willaschek 2018, 24–25; 64–70; Proops 2021, 423–24; Schafer 2023, 145–47; 200).

[38] 24:733; also see 2:418.

[39] 5:125; 5:144.

regulative principles but also doctrinal *Glaube* in some objects of traditional metaphysics. Either reading—and this is the essential point for current purposes—could be cashed out along epistemic consequentialist lines.[40]

Even if it's possible to read Kant as endorsing local epistemic consequentialism, however, I don't think this provides a response to the original worry. As I put it above, Kant arguably thinks we have warrant for optimistic beliefs about our own capacities, even when this is not adequately supported by evidence and the beliefs are in fact false. But if these beliefs are *known* to be false, we are rationally required to abandon them. He distinguishes between cases of ignorance and knowledge about whether our capacities are up to the task of accomplishing what we desire. We humans can become "aware" of evidence that defeats our presumption that we can accomplish a certain task, and thereby come to "know" the inadequacy of our capacities to bring about the ends in question.[41] If an agent with this knowledge continues to seek to bring about the ends in question, this agent is in contradiction with herself, and acts absurdly.[42]

To get a sense of this practical contradiction, it will help to first understand why it requires *explicit* knowledge of one's incapacity. There's an instructive contrast between humans and other animals, who are unable to get evidence that their capacities cannot match what they desire. That is, animals cannot acquire evidence that defeats their entitlement to mental representations on which they act.[43] Suppose an animal acts instinctively towards some end but only "in vain."[44] The animal cannot know that the end is unobtainable, so the animal cannot even be said to have a pointless desire for an end. For by definition, a pointless desire requires reflective knowledge *that* the desire is pointless. By contrast, humans can explicitly know that our ends are unobtainable, so we can have pointless desires for these ends. The crucial consequence is that if we do have this knowledge and still refuse to give up such pointless desires, we are in practical contradiction with ourselves.

To see why, start by considering a link Kant draws between desire and possible action. Desire aims at making the object of choice actual or real, and successful desire causes some object of desire to actually exist.[45] In turn, to choose an object is to determine the agent to an action that, at least in principle, can bring this object about. In aiming to bring about an object, desire also aims at action.[46] For this to be so, the agent must possess "a faculty for accomplishing" this end.[47] Setting aside the special case of moral action, the presence of this faculty means that it is really possible for the agent to cause the existence of the relevant end.[48] So my desires are appropriate only given a positive assessment that I really possess a faculty for accomplishing the desired end. If this assessment results in knowledge that *none* of my actions could bring about the end in question, then I also know that the end could not

---

[40] For versions of the stronger reading see Stang (2016, 280–96), Proops (2021, 4), Schafer (2023, 173), and Chignell (2024, 113). Chignell and Proops read these *Glaubens* as justified in part by their indispensability for good epistemic consequences. Some weaker readings, such as Briesen's (2013), also bring Kant close to epistemic consequentialism.

[41] 5:178. Also see 29:1013 on explicit "consciousness that" one cannot bring about an end.

[42] 5:178; 29:895.

[43] 29:895.

[44] 29:895.

[45] "The faculty of desire is the faculty to be, by means of one's representations, the cause of the objects of these representations" (6:211; also see 5:10; 5:178; 28:587).

[46] 6:381.

[47] 28:587.

[48] 5:20–21; 5:57–58; 6:213. By contrast, moral laws make no reference to what the agent can causally accomplish (5:21).

determine me to action. To continue to desire the end contradicts the aim of the faculty of desire, which is to cause the reality of its objects.

Kant's point is not that idle desires themselves are always self-contradictory. Instincts or passions are conative, but lack propositional structure. They cannot be literally self-contradictory. We see this when animals keep fruitlessly acting on instinct. Their instinct is not self-contradictory, just externally defeated and therefore pointless.

Instead, the contradiction is between a judgment that, in rational beings, is *presupposed by the desire*—namely that I have the ability to causally bring about what I desire—and my knowledge on other grounds that I lack the ability to bring about what I desire. While Kant sometimes calls this a contradiction *within* the faculty of desire, his considered view seems to be that there need not be contradiction in the faculty of desire "in itself"; the conflict arises from desire's being inevitably "deflected…by something else," namely knowledge that I cannot accomplish the desired end.[49]

As such, it is psychologically possible to knowingly desire a practically self-contradictory end, and even to act as if this desire were realizable—although this is to "behave unnaturally and absurdly."[50] Kant compares this to violating a logical law, for example by affirming a contradiction. The violation is excusable if I am unaware of a hidden contradiction. Once I recognize the contradiction, this innocence is lost and it becomes absurd to keep endorsing it. All the same, it may be psychologically possible to maintain this incoherent attitude.

The application of this practical contradiction to the OIC principle is as follows. Kant holds that if I rationally ought to φ, then on pain of practical contradiction, I must at least be capable of desiring to φ as an end. But if I know I can't φ, then I am not capable of desiring to φ as an end, so it won't be the case that I ought to φ. For 'φ,' fill in the three rational imperatives discussed in the previous section, namely Kant's regulative maxim, the *Necessary Injunction* that partly drives transcendental illusion, or the rational demand underlying the sublime. Now, an imperative both presents an end as desirable and prescribes end-directed action to bring it about.[51] If the end in question is unachievable, the imperative looks to be incoherent. As we've seen, this incoherence arises whenever we know an end is unattainable, regardless of whether the end is theoretical or moral. Once we *know* of the contradiction in *W(a)*, then, it will no longer be the case that we ought to follow these three rational imperatives.

3. The World as Noumenon

I now want to lay out a different response to the problem, which distinguishes between the troublesome contradictory concept and the concept that serves as the regulative idea of the world. Fully explaining this distinction will require delving into some of Kant's more recondite notions. Before getting into details, I give a less technical sketch of the origins of the contradictory concept.

The contradictory concept of the world, on my reading, derives from an apparently common-sense realist stance, on which what we can experience is the measure of things in themselves. From within this realist perspective, reflection on the world does not seem objectionably transcendent or metaphysical, or to require taking up an absolute point of view. Indeed, many of the claims driving contradiction seem at first to focus modestly on what can actually be given to us. This is illustrated by the traditional argument that because an actual infinity of past time could not be traversed, we must assume the world has a beginning in

---

[49] 20:230.

[50] 20:230; 29:895. Kant would agree with Locke and Reid that there are no restrictions on the objects of desire *if* this is just a point about descriptive psychology.

[51] 5:20.

time.[52] Part of the rationale for this argument is that an absolutely infinite magnitude is traversable only if, in principle, it is countable or enumerable. But complete enumeration is allegedly impossible, even for a sempiternal being. From the assumption that an actual infinite time cannot be experienced, the realist concludes that the spatiotemporal world is finite. Although this conclusion may seem epistemically modest, Kant argues that the realist tacitly attributes an absolute property to the spatiotemporal world, namely absolute temporal finitude. This entails contradiction. He diagnoses the realist's project as ultimately rooted in an immodest demand for the unconditioned. This demand is self-defeating because it can only be pursued from partial and conditioned spatiotemporal points of view. From within the point of view of common-sense realism, however, one sees neither that one is making a tacit demand for the unconditioned nor that one's spatiotemporal point of view is in fact conditioned.[53]

Now to bring in Kant's technical terms. Kant defines a noumenon as an *intelligibile* or object of the pure understanding, while a phenomenon is a *sensibile* or object of sense. *W(a)* is just the concept of a phenomenal world with absolute properties, I'll argue. The relevant regulative idea of the world, by contrast, is identical to the concept of a noumenal world with absolute properties. Then it turns out to be analytic, from the definitions of 'phenomenal' and 'noumenal,' that the regulative idea of the world and *W(a)* are different concepts. The contradiction Kant identifies in *W(a)* therefore has no bearing on the regulative idea of an infinite world.

3.1 The Phenomenon–Noumenon Distinction: A Closer Look

I next proceed to the textual evidence, and lay out six further points about the phenomenon-noumenon distinction that will clarify the two corresponding concepts of *world*. Unlike many commentators, I begin with definitions of *phenomenon* and *noumenon* that are neutral between Kant's transcendental idealism and dogmatic realism. Since part of the aim is to understand the concept *world* that Kant rejects as self-contradictory, we cannot presuppose his own transcendental idealist conclusions.

First, to say more on the difference between the concepts of *phenomenon* and *noumenon*. A noumenon is defined as that which can be cognized merely by the pure understanding and its pure concepts. It is an object that can be cognized independent from any contribution of sensibility.[54] A phenomenon is just what's left over, so any non-noumenon is a phenomenon. This means that the concepts of *noumenon* and *phenomenon* are extensionally disjoint. No phenomenon is a noumenon because as an object of sense, a phenomenon cannot be cognized through pure concepts alone. This does not exclude a role for pure concepts in cognizing phenomena. Phenomena might be cognized with the help of pure concepts, so long as they are not cognized through these concepts *alone*. Moreover, *phenomenon* and *noumenon* are defined in terms of mutually exclusive modes of *representation*. This definition does not entail that there exist corresponding disjoint classes of *fundamental entities*. One and the same fundamental entity could in principle have two modes of representation, one noumenal and the other phenomenal.

The second point is that since cognition of noumena is cognition by the understanding alone, it is a special type of a priori cognition. Some a priori cognition is not cognition of noumena, since it involves not only the pure understanding but also pure intuition. The objects of geometry are phenomena, for example, yet true propositions about geometrical objects have a priori proofs. Regarding geometrical objects as phenomena might seem to presuppose

---

[52] A432/B460.

[53] The tacit character of these absolute property ascriptions comes out clearly when Kant critically discusses a fallacious syllogism he attributes to dogmatic metaphysics (A497/B525ff.; for detailed reconstruction see Proops 2021, 257–61).

[54] A248–9; B306; 8:236; 28:560.

Kant's views on mathematics. He disagrees, asserting that traditional Platonism also treats geometrical objects as phenomena.[55] This raises an important interpretive possibility. The Antinomies lay out dogmatic propositions, purportedly on a priori grounds. But since not all a priori cognition is cognition of noumena, it need not be the case that the antinomial arguments concern noumena.

Third, the phenomenon–noumenon distinction crosscuts the difference between transcendental realism and transcendental idealism. Philosophers who deny noumena in favor of phenomena could understand the latter in a metaphysically noncommittal way. They might construe phenomena as whatever is manifest to intuition, leaving open whether what is manifest is also what is in itself.[56] Correspondingly, a philosopher might endorse the existence of noumena as objective correlates of a special type of cognition, without deciding whether these entities are things in themselves. So even if it turns out that noumena and things in themselves extensionally coincide, the concepts differ in intension. This possibility of drawing an intensional distinction between noumena and things in themselves is often overlooked. One reason is Kant's own tendency to equate things in themselves and noumena. This identification has become a convention among commentators.[57] However, the relevant texts might be read as recapitulating, not endorsing, a dogmatic assumption that noumena and things in themselves must coincide.[58] Another source of confusion may be that the main section of the first *Critique* devoted to phenomena and noumena takes for granted premises from Kant's own transcendental idealism. *Given* these additional premises, it may follow that phenomena are not things in themselves. But this does not follow from the definition of *phenomenon*.[59] And that is crucial for the opposed arguments of the Antinomies, which stem from the perspective of a transcendental realist. They exploit the definitional distinction between phenomena and noumena, but they *cannot* assume what transcendental idealism says about noumena and phenomena, since the Antinomies are supposed to indirectly *prove* transcendental idealism.[60]

This point can be extended with a fourth observation, namely that phenomena are invoked not just in the natural attitude of science and everyday life, but also in dogmatic metaphysics. Kant sees Aristotle as building intuitions of space and time into his categories. This entails that spatiotemporal intuition is an element of all cognition whatsoever, and that existence in a world is necessarily spatial. This "cosmological materialism," as Kant calls it, is immodestly dogmatic.[61] Aristotle rules out objects of cognition that are independent of intuition: he leaves no room for noumena.

---

[55] Kant depicts Plato and Pythagoras as concluding from mathematics that humans "possess intuitions a priori," and then reasoning that because our understanding is "discursive," these intuitions must be rooted in something radically outside us—Platonic Forms or "an understanding that rules over" nature (8:391–92). In other words, Platonists take our representation of Forms as phenomenal because this is the only way to use Forms to explain mathematical truth.

[56] See further Ameriks (2003, 109–10), Allais (2015, 117–24), and Jauernig (2021, 338–44).

[57] See among many examples 4:312; 4:316; 20:308; Willaschek (2018, 142; 152; 245).

[58] See Hogan (2021, 398–99; 412); Jauernig (2021, 338).

[59] The contrast between phenomena and things in themselves at B306, for example, invokes premises from the Transcendental Aesthetic about the subjectivity of our "way" of intuiting. Kant also needs his thesis that concepts without intuitive content are empty in order to conclude that we have no determinate theoretical cognition of noumena. Kant only seems to fully arrive at this diagnosis of noumena in the *Critique*'s second edition, and may neglect to revise relevant passages, adding to the obscurity of his discussion (Allison 2004, 63; Allais 2015, 61–65; Jauernig 2021, 341n.71).

[60] See the so-called indirect argument for transcendental idealism at A506–7/B534–35.

[61] 6:128n.; also see A81/B107; 8:393; 28:9. As Kant notes (2:76), his contemporary Christian Crusius defended cosmological materialism in this sense.

Fifth, from the perspective of traditional realism, noumena collectively make up an intelligible world that is distinct from the sensible world of phenomena.[62] Phenomena, by definition, are non-noumenal objects. So from this definition, it follows that the concept of a phenomenal world is disjoint with the concept of a noumenal world. Since Kant is not a traditional realist, he does not think this point settles all metaphysical questions about how phenomena and noumena relate. He does, however, have resources to block the following tempting line of thought about noumenal and phenomenal worlds. Since by definition (1) the noumenal world contains everything thinkable through mere concepts, one might suppose that if (2) we can think in this way about all possible phenomena, then (C) any possible phenomenal world would necessarily be a proper part of the noumenal world. Kant denies premise (2). *No* phenomenon can be thought through "mere concepts, i.e., a concept of a noumenon."[63]

This brings me to the sixth and final point. Given transcendental idealism, one might expect Kant to say that there is not even a logically consistent concept of a noumenal world. But instead, he is explicit that *W(a)*, which is self-contradictory and drives the Antinomies, is the concept of a world of *phenomena*, not noumena. The self-contradictory concept *W(a)* does not overstep "the object, namely appearances, *in kind*," just because *W(a)* has to do "*only* with the sensible world (*not* with noumena)."[64] Elsewhere, he states that *W(a)* "concerns nothing other than the exposition of appearances," and that to consider the world "by its beginning or infinite duration" is to consider it "as phenomenon," rather than as noumenon.[65] Yet another passage explains that the Antinomies could "only" be *resolved* through the "cosmological idea of an intelligible world"—so this idea of a noumenal world cannot result in antinomial contradictions.[66] Since *W(a)* is a concept of the phenomenal world, it is at least extensionally distinct from any concept of the noumenal world, which we can call *W(n)*. Since noumenal and phenomenal representations are disjoint, *W(a)* cannot be a concept of a noumenal object.

3.2 Consistent Infinity

To give these points concrete application, consider again the conflicting arguments about the eternity of the world in the first Antinomy. These arguments appeal to what we can and cannot experience. The thesis argument, for example, assumes for reductio that the world is eternal. But this, the argument goes, entails a completed infinite series of times leading up to any given moment. And, allegedly, it's analytic that an infinite series cannot be completed by proceeding stepwise through times, since each time is finite.[67] So there cannot be such a series, and the world cannot be eternal.

Now, it might seem that this argument presupposes Kant's own transcendental idealist position. How, if not from previous acceptance of idealism, could the limitations of successive synthesis settle questions about the *world*?

The question misconstrues the dialectical status of these arguments. They are presented from the perspective of a transcendental realist who takes the phenomena, or what can be

---

[62] 4:314.

[63] A285/B342; also see 28:560.

[64] A420/B447; emphases added.

[65] A416/B443; 20:238. These passages confirm the need to distinguish between noumena and things in themselves. In employing *W(a)*, transcendental realists equate spatiotemporal phenomena with things in themselves. Since the concepts of phenomenon and noumenon are extensionally disjoint, realists cannot thereby equate phenomena with noumena. Therefore, noumena are not things in themselves.

[66] 5:133.

[67] A426/B454; A740/B768. The realist apparently considers the series as committed to an infinite *number* of times. That would violate the Aristotelian ban on traversing an infinite number of *X*s.

experienced with the help of the senses, to suffice for conclusions about the world in general, as it is in itself. Kant describes the antithesis argument as "in accordance with the *common and dogmatic way of representing*…[letting] the world of *sense* count as a thing whose totality is given in itself prior to any regress."[68] He might well agree that we can't experience the completion of an infinite series, and so cannot experience that the phenomenal world has the property of absolute temporal infinity. He does not conclude that the phenomenal world has the property of absolute temporal finitude. The conflict in the first Antinomy arises from attempts to ascribe absolute temporal and spatial properties to the phenomenal world. The Antinomy need not refer to the noumenal world at all.

Granting that *W(a)* differs from *W(n)*, a worry remains. If the concepts are sufficiently similar in content, the contradiction in the former concept might also infect *W(n)*. I now want to defuse this worry by looking more closely at the origins of the contradiction in *W(a)*.

What is at fault, Kant thinks, is the attempt to combine the concept of a phenomenal world with the rational concept of the unconditioned, and more specifically from attributing absolute properties to the phenomenal world. The concept of the unconditioned is roughly the concept of an ungrounded ground. This "necessary rational" concept of the unconditioned is logically consistent.[69] To have determinate content, however, the concept of the unconditioned needs to be further specified in terms of more particular relations. For example: an unconditioned proper part is mereologically simple, or an unconditioned efficient cause is uncaused.

Contradictions lurk in these putative specifications of the concept of the unconditioned. Kant uses an analogy with chemical reactions to bring this out. The Antinomies confirm that "things as appearances" and "things in themselves" are "heterogeneous" because they *react differently* to "unison with the necessary rational idea of the unconditioned."[70] What is analogous to two different chemical reactions is the presence or absence of contradiction. That is, an attempt to combine the concept of the unconditioned with the concept of things as appearances yields a contradictory concept.

Contradiction arises because the concept of the phenomenal world essentially represents entities as standing in spatial and temporal relations. Kant takes it to be a conceptual truth that all spatial and temporal relations are conditioned. He holds further that to combine the concept of the phenomenal world with the concept of the unconditioned is already to commit to the possibility of unconditioned spatial and temporal relations. But spatial and temporal relations are essentially conditioned. Unconditioned spatial and temporal relations are self-contradictory, hence impossible.

More precisely, *W(a)* is what Kant calls an *inferred concept*.[71] An inferred concept results if the premises in a deductive argument ascribe properties, say *F* and *G*, to some individual *x*. Then since *x* is *F* and *x* is *G*, it can be inferred that *x* is *F and G*. Furthermore, the premises are taken as committing to the existence of *x*. The conclusion is that there is at least one individual that instantiates the concept of *being F and G*.

---

[68] A521/B549n.; my emphases. This argument invokes the principle that the world's properties must be "determinate," specifically that the whole phenomenal world has a well-defined "place in" space and time (A521/B549n.; and compare 2:388–89; 2:415). See further Ameriks (2003, 110).

[69] Bxx. The general concept of the unconditioned cannot be self-contradictory because it turns out to have practical reality (Bxviii; A417/B445; 4:353–57; Hogan 2021, 420–30). Kant seems to assume any unconditioned must be thought as a noumenon. This is not obvious: some might say pleasure is an unconditioned good, but no noumenon.

[70] Bxxi. Also see A407/B433 and A417–18/B445(n.).

[71] A310–11/B366–68. I follow Willaschek's (2018, 175–77) helpful discussion of inferred concepts. As he observes, Kant indicates that the relevant inferences can be tacit and involuntary (208).

Here, an inferred concept is formed through an unsound inference. Although the inference to the existence of an individual instantiating *W(a)* is unsound, it contains true premises predicating spatiotemporal properties of phenomenal objects. For example, Kant accepts that the world contains bounded spatial regions. Along with these acceptable premises, there is at least one false premise in the argument. The false premise ascribes some unconditioned property to the phenomenal world, for example that it contains a complete series of spatial regions.[72] This attribution confuses noumenal and phenomenal world-concepts.[73] The false inferred conclusion is that there exists some individual—*the world*—that has not only finite phenomenal properties, but also unconditioned phenomenal properties. A property is unconditioned in the relevant sense if it is "a presupposition that presupposes nothing further."[74] A region of space is unconditioned if it encompasses all actual spatial parts while not being a part of any larger whole. Kant takes himself to have shown that "everything in space and time is conditioned (internally)," so it is self-contradictory to ascribe the relevant unconditioned properties to the "sensory world" or its parts.[75]

To see what Kant means in saying that things in space and time are internally conditioned, consider a stone. As a determinate concrete whole, the stone is composed out of parts through aggregation. Insofar as it has a magnitude, the stone essentially has asymmetric conditions or presuppositions: its parts and their aggregation into a specific arrangement. Less intuitively, Kant also holds that the space the stone occupies is conditioned by the surrounding space. Only in this way can the stone occupy a determinate space. This should shed light on why regarding the phenomenal world as having unconditioned quantitative properties, as in the concept *W(a)*, is inconsistent. It is inconsistent to ascribe an unconditioned property of type *F* to an entity for which, necessarily, all properties of type *F* are conditioned.

The contradiction in *W(a)* thus arises not from the sheer attempt to think the unconditioned, but from an attempt to think unconditioned *phenomenal* properties. We can consistently think of the world as "absolute whole" or "totality" just in case the world is "thought as a noumenon."[76] Lectures from 1782–83 state that there is no contradiction in the concept of an intelligible world made up of an "infinite multitude" of noumena.[77] To "ascribe limits to" the world "qua noumenon" would simply be "false": a true representation of the

---

[72] Some might object that the Thesis of the second Antinomy defends *simple substances*, and that these are noumena. This is a larger topic that I can cover here, but briefly: the Second Antinomy deals strictly with extended composites *in space*, or what Kant elsewhere calls phenomenal substances (A441/B469; 29:825–27; Proops 2021, 238). The Thesis argument runs from spatial phenomena to conclusions about simple spatial substances. My hunch is that Kant regards dogmatic monadists as not entirely clear on whether their monads are noumena or phenomena (A439/B467; 4:237–38). He sometimes depicts Leibniz's monads not as noumena, but as phenomena wrongly *taken for* noumena (A264/B320). On this reading, Leibniz thinks monads are in space and time, but are not directly observed by our limited senses. Other passages suggest it would be more "proper" to regard monads as noumena (8:248–49), and that this is the most charitable reading of Leibniz (4:508). However, Kant is explicit that the second Antinomy is *not* concerned with any such noumenal monads, but with spatiotemporal simples (A442/B470).

[73] In turn, *this* confusion is itself due to prior erroneous inferences (29:849; A516/B544). For details, see Willaschek (2018, 204–208).

[74] A323/B380.

[75] 20:328; 20:288. See also A483/B511 ("in space and time…the whole…is always only comparative"); B202; 4:342; 5:104.

[76] 20:328. Also see 29:834: "there is with the noumena a maximum (greatest)."

[77] 29:837; A793/821 also commits to the consistent thought of a whole world with "unconditioned" "infinity." Now, Kant holds that *infinite number* is a contradictory concept, but this need not conflict with his endorsement of an "infinite multitude," since a multitude need not be a number. I take the putative contradiction to arise from the assumption that numbers must be countable, plus the uncountability of the infinite (29:835; 29:841; 29:994). But he allows that infinities need not be numbers (A430/B458).

world qua noumenon must represent the world as unlimited.[78] So by the principle of excluded middle, to say that any noumenal world is essentially unlimited or infinite is not just consistent, but *true*.

The point is confirmed when the lectures ask whether there is a contradiction in the concept of a world that's both composite and absolutely complete. The answer is 'no':

> The world is no relative whole, but rather an absolute whole in the metaphysical sense…We cannot experience the totality of the world…But then are these transcendental…ideas…not mere phantoms of the brain? No. It is necessary for reason to bring all of its concepts to completion and therefore also to make complete the absolute composite, for nature brings with it the projecting of general rules, therefore it can stand for nothing incomplete.[79]

So the concept of the world as a complete composite is not contradictory or a mere phantom. To the contrary, it is rationally necessary to bring concepts of magnitude to completion through a noumenal conception of the world. To do so is to represent a complete, absolute composite.

A nearby passage reaffirms the distinction between *W(n)* and a self-contradictory concept of an absolute phenomenal world:

> we can consider the world in general as noumenon…I abstract here from the manner in which such a whole of substances supposedly can be intuited, and find no contradiction in this idea of reason. I also find no contradiction in this absolute whole not occurring in appearances or in the sensible world.[80]

The concept of a phenomenal world turns out to be the concept of a world that is essentially spatial and temporal. There is no way to intuit absolute size properties of a phenomenal world, since these properties would have to be spatiotemporally intuited and intuited properties cannot be absolute. By contrast, the concept of a noumenal world abstracts from the way in which such a world would be intuited.[81] So there is no contradiction in the concept of a noumenal world as endowed with absolute magnitude properties, such as infinity.

3.3. The Noumenal World and Inquiry

We can now return to the consistency problem with Kant's regulative concept of the world. I propose that the regulative concept is the logically consistent concept of a noumenal world with absolute properties, rather than the self-contradictory concept of a *phenomenal* world with absolute properties.

Notes on his metaphysics lectures report Kant saying exactly this. Just because the noumenal concept of the world is logically consistent and not a mere phantom of the brain, this "concept of the whole cosmos" can be "a concept of reason" that guides inquiry: it can "easily" be thought, even if its object cannot be given concretely.[82] The concept of a noumenal world may lack a referent, but it is not self-contradictory.

For further support, consider published passages where Kant identifies this regulative idea with the consistent concept of a noumenal world:

> *All…noumena,* together with their aggregate—*an intelligible world*—are nothing but representations of a problem, whose object is *in itself perfectly possible*, but whose solution, given the nature of our understanding, is completely impossible.

---

[78] 20:328. In modern parlance: if a noumenal world exists, then necessarily it is unlimited.

[79] 29:851–852.

[80] 29:853–54.

[81] 29:853.

[82] 29:851; translation modified.

> We should…think for ourselves an immaterial being, *an intelligible world*, and a highest of all beings (*all noumena*), because only in these things, as things in themselves, does reason find completion.
>
> If those concepts should yield, not constitutive, but merely regulative principles of the use of reason (*as is always the case with the idea of a noumenon*), they can also, as merely logical functions for the concepts of things whose possibility is unprovable, have a use for reason that is indispensable to it from a practical viewpoint, because they would then be valid, not as objective grounds of the possibility of noumena, but as subjective principles (of the theoretical or practical use of reason) in regard to phenomena.[83]

So noumena are intentional objects of consistent regulative ideas. Commitment to noumena in what Kant calls a negative sense suffices for this. He accepts that ideas of noumena are logically consistent, but denies that the real possibility of the objects these ideas represent can be established on theoretical grounds. This is also how he describes the intentional objects of theoretical ideas of pure reason.[84] By contrast, a noumenon in a positive sense would be an object of positive cognition. We do not have positive theoretical cognition of noumena, so there are no noumena in the positive sense.

One might ask how a representation of a noumenal world could guide inquiry into the phenomenal world. The question can be sharpened by considering a condition Kant places on part–whole relations: they only obtain between relata that are both of the same general kind, or homogeneous. A spatial region can have smaller spatial regions as proper parts, for example, but cannot have something non-spatial as a part. If *W(n)* represents an infinite whole, this whole is not spatial or temporal, because it is characterized by pure concepts alone. Parthood relations only obtain between relata of the same general kind, but non-spatial wholes are *not* of the same general kind as spatial parts. So an infinite totality of noumena could not stand in part–whole relations with any spatial region or temporal interval. This seems to preclude any role for *W(n)* in guiding inquiry in the phenomenal world.[85]

The best response, I think, is to insist that the regulative force of *W(n)* does not depend on possible parthood relations between phenomena and noumena. The intentional objects of regulative ideas need not be represented as standing in token conditioning relations to token phenomena in order to play their role. Here is Kant's summary of the regulative purport of the idea of the world, which also cautions against invoking the noumenal world to explain particular phenomena:

> We have to pursue the conditions of the inner as well as the outer appearances of nature through an investigation that will nowhere be completed, as if nature were infinite in itself and without a first or supreme member—although, without denying, outside of all appearances, the merely intelligible first grounds for [the appearances of nature], *we may never bring these grounds into connection with explanations of nature, because we are not acquainted with them at all*.[86]

In the first part of this sentence, Kant seems to head straight for incoherence. The regulative concept of nature as infinite in itself, that is, of nature as noumenon, prescribes a pursuit of conditions in the phenomenal world. This could be read as saying that nature as a noumenon

---

[83] 4:316; 4:354; 8:225. Emphases added.

[84] Compare A253–5/B308–10 with A325–27/B381–83 and A673–75/B701–3. And see readings on which Kant is something like a fictionalist about both noumena and the intentional objects of ideas of theoretical reason, though *not* a fictionalist about things in themselves (Allais 2015, 65–70; Jauernig 2021, 336–37).

[85] On homogeneity, see e.g. 4:343. For the worry that representations of noumena cannot guide inquiry into heterogeneous phenomena, see Jauernig (2021, 334).

[86] A672/B700; my emphasis. Also see A483–84/B511–512; A602/B630; 4:353.

is represented as conditioning particular phenomena. But in the last two clauses, Kant strictly rules out invoking noumena as grounds or conditions for explanations of particular appearances in nature. The representational content of the relevant idea is just that of a complete noumenal world. Given Kant's homogeneity requirement, we cannot assert that the phenomenal world, or any phenomenon in it, is a spatial or temporal part of the noumenal world. Nor is the phenomenal world composed out of noumena as parts. All explanations of nature under consideration here are explanations of phenomena. If we wish to explain phenomena in terms either of their parts or of larger wholes that contain them, then we must appeal to further phenomena.

The last quoted passage suggests how the idea can have regulative force, even if it does not explain particular phenomena. If we confined our attention to the phenomenal world, we would have no representation of "an investigation that will nowhere be completed." In the material world, we just discover determinate spaces and times. These determinate spaces and times are always finite.[87] Based on this evidence, it is an open question whether further progress in our investigations will remain possible. Our idea of the noumenal world, by contrast, represents a world with a completed infinite extent. So it represents a world that remains inexhaustible however far inquiry might proceed.

Further textual support for this proposal can be gleaned from an otherwise puzzling passage on the regulative idea of the world. Natural objects, Kant states, come in just two kinds: thinking things and corporeal things. He goes on to assert that the regulative idea of the world does not apply to *either* type of natural thing. Instead, "for pure reason" in its regulative use, "there is nothing left…except nature in general."[88] By the idea of nature in general, we can understand the idea of a complete noumenal world in abstraction from token explanatory or grounding relations to phenomenal objects. The same passage stresses that the relevant world-concept depicts a series that is "in itself infinite."[89] While infinite wholes themselves "can never come about fully" in the phenomenal world, a consistent concept of them is needed to articulate a "rule" for theoretical inquiry, namely to always "proceed *in indefinitum*."[90]

To sum up, inquiry requires representing a completed actual infinity of objects. We can never represent this on the basis of mere phenomena. A representation of a noumenal world, by contrast, does allow for such a representation. Despite Kant's charge that the traditional metaphysical concept of an infinite world is inconsistent, a consistent rule for inquiry remains available to him.

3.4 Two Problems Resolved

Finally, let me sketch how this solution can also dissolve the other problems I considered at the start.

To review the second problem, Kant's error theory for illusions of dogmatic metaphysics links them to legitimate assumptions, particularly the *Necessary Injunction* discussed above. But if these supposedly legitimate assumptions are themselves based on a self-contradictory concept of the world, the error theory will collapse. My interpretation addresses this problem because instead of the inconsistent *W(a)*, the consistent representation

---

[87] Space and time as mere forms of intuition are infinite, but they are indeterminate and non-material (see references in footnote 6 above).

[88] A684–85/B712–13.

[89] A685/B713.

[90] A685/B713. For more on the normativity of regulative principles, see Allison (2004), Willaschek (2018), Proops (2021), and Kraus (2025). If my argument here succeeds, then *any* reading on which the idea of the world is the idea of a noumenon has the decisive advantage of avoiding contradiction. Defending my own interpretation of how regulative guidance works in detail is beyond the scope of this article.

of an intelligible world *W(n)* underwrites *Necessary Injunction*. So *Necessary Injunction* itself is consistent, and does not on its own lead to the inconsistencies of dogmatic metaphysics. Problems only arise with attempts to follow the injunction in the case of phenomena, when series of phenomenal conditions are taken as unconditioned.

The third and final problem concerned the mathematical sublime. The experience of the sublime, springing from infinite magnitude properties of the world, seems inextricably linked to an inconsistent concept of the world as possessing unconditioned spatiotemporal properties. Kant says quite plainly that the *only* way to avoid this threat of contradiction is to construe the concept of the world as the concept of a noumenon:

> Even to be able to think the given infinite *without contradiction* requires a faculty in the human mind that is itself supersensible. For it is *only by means of this and its idea of a noumenon*, which itself admits of no intuition though it is presupposed as the substratum of the intuition of the world as mere appearance*, that the infinite of the sensible world is completely comprehended in the pure intellectual estimation of magnitude under a concept*….A faculty for being able to think the infinite of supersensible intuition as given (in its intelligible substratum) surpasses any standard of sensibility…not…from a theoretical point of view, on behalf of the faculty of cognition, but still as an enlargement of the mind which feels itself empowered to overstep the limits of sensibility from another (practical) point of view.[91]

This passage affirms that there is a purely intellectual way to think of an infinite magnitude, without any use of intuition, though this does not yield theoretical cognition. Kant contrasts the consistent idea of a noumenal world with the inconsistent concept of a spatiotemporal world that is infinite in extent (specifically, the inconsistent prospect of comprehending such a spatiotemporal world as a whole). However, the end of the passage reminds us that no determinate theoretical cognition of noumena follows from this.

4. Conclusion

Kant's references to noumena may at first seem obscure, or at best to obliquely make points better expressed in terms of things in themselves. But if the arguments of this article succeed, then following the details of the theory of noumena—where, among other things, this means appreciating their definitional distinctness from things in themselves—is needed in order to understand how his norms of inquiry can avoid inconsistency.[92]

---

[91] 5:254–55; my emphases. So I disagree with Winegar's (2022, 655–57) reading, where infinite "space and past time…as given" can consistently be thought (5:254). For Kant directly calls such a thought "self-contradictory," and attributes it to mere "common reason," prone to err in metaphysical matters (5:256; 5:255; cf. A461/B489). What may be going on here is that space and past time as *potential* infinities force common reason to some inchoate idea of nature as infinite. However, the *only* way to *consistently* combine that idea with a rational demand for "comprehension" (hence actual, completed infinity) is to throw away the sensible ladder and think infinite, "supersensible" nature (5:255). On rational comprehension, see again Schafer (2023, 119–148).

[92] I am grateful to Juan Carlos González for helpful comments on an earlier draft. For stimulating discussions of related issues, thanks also to Katharina Kraus, Nate Lauffer, Roy Sorensen, Lorenzo Spagnesi, and participants in Andrew Chignell's "Kantian Epistemologies" seminar at Princeton: Larissa Berger, Claudi Brink, Pirachula Chulanon, Eliza Little, and Jake McNulty.

References

All references to Kant's works are to volume and page number of the *Akademie* edition (*Kants gesammelte Schriften*), except for references to the *Critique of Pure Reason*, which is cited according to the standard A/B pagination.

Allais, L. (2015). *Manifest Reality*. Oxford: Oxford University Press.

Allison, H. E. (2004). *Kant's Transcendental Idealism*. New Haven: Yale University Press.

Ameriks, K. (2003). *Interpreting Kant's Critiques*. Oxford: Oxford University Press.

Briesen, J. (2013). Is Kant (W)right? On Kant's Regulative Ideas and Wright's Entitlements. *Kant Yearbook*, 5, 1–32.

Chignell, A. (2024). Kant on Theoretical (or 'Doctrinal') Belief and the Status of Metaphysics. *Studi Kantiani*, XXXVII, 105–126.

Englert, A. T. (2017). Dutifully Wishing: Kant's Re-evaluation of a Strange Species of Desire. *Kantian Review*, 22(3), 373–394.

Hadisi, R. (2022). Kant's Account of Epistemic Normativity. *Archiv für Geschichte der Philosophie*. doi:10.1515/agph-2021-0058

Hogan, D. (2021). Handedness, Idealism, and Freedom. *The Philosophical Review*, *130*(3), 385–449.

Jauernig, A. (2021). *The World According to Kant*. Oxford: Oxford University Press.

Kant, I. (1900–). *Kants gesammelte Schriften* (Preussische Akademie der Wissenschaften, Eds.). Berlin: de Gruyter.

Kraus, K. T. (2025). *Kant's Ideas of Reason*. Cambridge: Cambridge University Press.

Leech, J. (2017). The Normativity of Kant's Logical Laws. *History of Philosophy Quarterly*, 34(4), 351–370.

Lu-Adler, H. (2017). Kant and the Normativity of Logic. *European Journal of Philosophy*, 25(2), 207–230.

Proops, I. (2021). *The Fiery Test of Critique*. Oxford: Oxford University Press.

Schafer, K. (2023). *Kant's Reason*. Oxford: Oxford University Press.

Stang, N. F. (2016). *Kant's Modal Metaphysics*. Oxford: Oxford University Press.

Stern, R. (2004). Does 'Ought' Imply 'Can'? And Did Kant Think it Does? *Utilitas*, 16(1), 42–51.

Vaihinger, H. (1924). *The Philosophy of 'As If'* (C. K. Ogden, Tr.). New York: Harcourt, Brace, & Co.

Willaschek, M. (2018). *Kant on the Sources of Metaphysics*. Cambridge: Cambridge University Press.

Winegar, R. (2022). Kant's Three Conceptions of Infinite Space. *Journal of the History of Philosophy*, *60*(4), 635–659.