

The Many Faces of Transcendental Realism: Willaschek on Kant's Dialectic

ANDREW CHIGNELL

Princeton University

Email: chignell@princeton.edu

Abstract

After providing a brief overview of Marcus Willaschek's *Kant on the Sources of Metaphysics*, I critically reconstruct his account of 'transcendental realism' and the role that it plays in the dramatic narrative of the *Critique of Pure Reason*. I then lay out in detail how Willaschek generates and evaluates various versions of transcendental realism and raise some concerns about each. Next, I look at precisely how Willaschek's Kant thinks we can avoid applying the 'supreme' dialectical principle (for every conditioned there is a totality of conditions which is unconditioned) to the domain of appearances. Finally, I call into question Willaschek's efforts to appropriate the lessons of the Transcendental Dialectic without following Kant into transcendental idealism.

Keywords: Kant, *Critique of Pure Reason*, Transcendental Dialectic, transcendental idealism, metaphysics, critique

1. Overview of the Book

Marcus Willaschek's elegant, readable book (*Kant on the Sources of Metaphysics: The Dialectic of Pure Reason*, Cambridge University Press 2018) combines an innovative interpretation of Kant's overall approach in the Transcendental Dialectic (Part I) with a series of focused accounts of the more familiar arguments against speculative metaphysics (Part II). I will focus on Part I here: the overarching framework Willaschek provides is important because it helps us decipher some key mysteries in Kant's account of reason and its basic needs and interests, as well as of its metaphysical illusions and mistakes.

Highlights of this part of the book include:

- An overall ‘template’ for thinking about Kant’s arguments in the Dialectic according to which reason naturally starts with various ‘logical’ principles and then transitions to analogous ‘real’ principles, ultimately misusing them (in the absence of critique) by going beyond regulative, heuristic purposes to illegitimate constitutive system building.
- An account of what Kant really thought about the rationalist principle of sufficient reason in the critical period.
- An account of the ‘natural argument’ that (for Kant) we are all tempted to follow with respect to God, of why it is flawed and of how it relates to the classic theological proofs.
- An argument for the so-called ‘radical reading’ of Kant’s claim that transcendental ideas are ‘empty’.
- A four-level model of how Kant’s argument proceeds in each chapter of the Dialectic.

All this delivers two main results.

The Central Exegetical Result

Willaschek defends what he calls the ‘Rational Sources Account’ of Kant’s views about speculative metaphysics. That account, roughly, says that discursive reason raises a metaphysical question, is driven by its own ‘need’ or ‘nature’ to answer it and then iterate, and then answer again, and then reiterate, and so on – until it achieves completeness. Moreover, the principles it uses to do this are principles arising from ‘universal human reason’: they ‘belong to rational thinking as such’ (p. 5). Kant’s great discovery is that these principles ‘work fine within the realm of experience’ (p. 4), but when we allow ourselves to follow the train of reasoning all the way to a non-empirical unconditioned, we find ourselves falling into ‘fallacies and contradiction’ (p. 5).

The Central Philosophical Thesis

Willaschek allows that Kant ‘indeed discovered a source of metaphysical thinking that lies in reason itself’. The *discursive, iterative* and *completeness-seeking* character of reason explains why we naturally and (in some sense) rationally move from, in Kant’s terms, the conditioned to the totality of conditions and ultimately to the unconditioned (p. 6). Even if Kant’s prescription for the problem (i.e. adopt transcendental idealism!) is not viable for many philosophers, Willaschek argues at length (in chapter 9) that Kant’s diagnosis is useful and points to other ways of thinking about how we can avoid it.

This brief summary indicates how much fascinating interpretative and philosophical material there is in the book. It is well worth reading by anyone interested in Kant's account of reason and his approach to general and special metaphysics. Here I will have to circumscribe my goals and discuss just a few key issues.

My main goal (in sections 2 and 3) is to consider Willaschek's account of 'transcendental realism' and the role that it plays in the dramatic narrative of the *Dialectic*. I will lay out in some detail how he generates and evaluates the various versions of transcendental realism and then raise some concerns about each. In section 4, I will consider how, exactly, Willaschek's Kant thinks we can avoid applying a key dialectical principle (for every conditioned there is a totality of conditions which is unconditioned) to the domain of appearances. Finally (in section 5), I will problematize Willaschek's efforts in chapter 9 to appropriate the lessons of the *Dialectic* without following Kant into transcendental idealism.

2. The Origins of Transcendental Realism

One of the more controversial but also central components of Willaschek's book is his account of what 'transcendental realism' is and how it functions in the *Dialectic*. He first introduces it as the claim that

the structure of reality corresponds to that of rational thought or, more generally, the subjective conditions of thinking rationally about objects are conditions of the object being thought about. (p. 9)

Later he admits that this is a view more prominent in the rationalist tradition than in the empiricist tradition, and explicitly locates versions of it in Plato, Descartes, Spinoza, Leibniz and others (pp. 144–8). For the time being, let us just call it 'TR_{rationalism}':

TR_{rationalism}: The structure of reality corresponds to that of rational thought.

Chapter 5 is where Willaschek discusses transcendental realism (TR) at great length, introducing it with this quotation about its alleged opposite, transcendental idealism (TI):

everything intuited in space or in time, hence all objects of an experience possible for us, are nothing but appearances, i.e. mere representations which, as they are represented, as extended

beings or series of alterations, have outside our thoughts no existence grounded in itself. This doctrine I call **transcendental idealism**. The realist, in the transcendental sense, makes these modifications of our sensibility into things subsisting in themselves, and hence makes mere representations into things in themselves. (A490–1/B518–19)

A natural reading of this passage says that transcendental idealism is the doctrine that spatiotemporal features are not fundamentally real, and thus that the objects we experience as having spatiotemporal features are, as such, ‘nothing but appearances’ that are somehow dependent (for their features and/or existence) on minds. Willaschek endorses this reading, and sums it up in the following thesis:

Transcendental Idealism (TI): ‘Empirical objects are not things-in-themselves’ (p. 138). They are not things in themselves because ‘all the properties of [these empirical] objects we can cognize (“appearances”) depend (in some appropriate sense) on the possibility of being represented by finite rational beings like us’ (p. 247).

What about transcendental realism (TR)? In the A490/B518 passage just quoted, it looks like the ‘realist in the transcendental sense’ simply denies TI: her primary commitment is to the claim that objects with spatiotemporal features are not ideal but rather *real* in some fundamental, mind-independent sense – i.e. they are ‘things in themselves’. Thus:

TR_{things}: ‘Empirical objects are things-in-themselves’ (p. 140).

What I do not see in this or any passage is the claim that TR_{things} ultimately boils down to or follows from TR_{rationalism}. For TR_{rationalism} says that the structure of reality corresponds to the structure of rational *thought*. But TR_{things} is articulated in the passage as the claim that ‘the modifications of our sensibility’ (i.e. spatiotemporal appearances or the spatiotemporal features that objects appear to have) are (or attach to) things in themselves. That is in effect to say that (the structure of) reality corresponds to the way our *experience* (and specifically our *sensible* experience) represents it, rather than the way we rationally *think* about it.

So one thing that almost any reader will wonder is: why does Willaschek think TR_{rationalism} counts as an adequate statement of what Kant seems to say in this passage about ‘realism in a transcendental sense’? More simply put: what is the connection between TR_{rationalism} and TR_{things}?

3. The Evolution of Transcendental Realism

The versions of TR proliferate as we go along. Willaschek next argues that things in themselves are identical to 'noumena' (things of thought), and that because noumena can be conceived both negatively and positively there is an ambiguity when we substitute 'noumena' for 'things' in TR_{things} (p. 142):

TR_{neg}: Empirical objects are noumena in the negative sense.

TR_{pos}: Empirical objects are noumena in the positive sense.

Willaschek mentions that there is controversy over how to interpret the idea of noumena in the negative sense. In a 1998 publication he argued for a reading according to which a negative noumenon is just an empirical object with its sensible features abstracted away. In other words, a noumenon *is* an empirical object but 'not considered as an object of the senses'.

If that view is coherent, then it would make sense of TR_{neg}. But in the present context Willaschek (following Kant) wants TR_{neg} to turn out false, and so he instead construes it as the claim that empirical objects are 'non-sensible (not accessible to our senses)' (p. 142). This makes TR_{neg} a clear non-starter, and Willaschek sets it aside. It is left unclear whether this is a change in his view regarding what a noumenon in the negative sense is.

In any case, once we agree that TR_{neg} is obviously false and thus not at all tempting to reason or common sense, that leaves TR_{pos}. Sure enough, Willaschek says that 'what Kant means by "transcendental realism" can only be TR_{pos}' (p. 143). But what exactly *is* a noumenon in the positive sense? Willaschek writes:

Kant's main point in introducing the concept of a noumenon in the positive sense, apart from distinguishing it from the concept of a noumenon in the negative sense, is to highlight that we can consistently *think* of a class of objects that finite, sensible beings like us cannot *cognize* (B146). These are non-sensible objects or noumena in the positive sense. (pp. 141–2)

The 'positive' part of the conception involves conceiving of them as 'objects of a hypothetical non-sensible or intelligible intuition – a kind of intuition Kant typically attributes to God' (p. 141).

So now we have arrived at the view that, for Kant, TR is the ‘weighty metaphysical claim’ (p. 143) that empirical objects are ‘non-sensible objects or noumena in the positive sense’ which ‘beings like us cannot cognize’ and which ‘can be cognized only by an infinite being (God)’ (pp. 141–2).

But if TR is supposed to be at all plausible, much less tempting to reason and common sense (as Kant says that it is), it is hard to see how this can be right. How can transcendental realism be the claim – which seems like another obvious non-starter – that *empirical* objects are ‘non-sensible objects’, i.e. objects that we ‘cannot cognize’ and that ‘can be cognized only by an infinite being (God)’? Willaschek does not say anything about abstracting from sensible properties, or empirical objects ‘considered as’ non-spatiotemporal objects: it is just the straight-up claim that empirical objects are non-sensible objects that only God can cognize. But the assumption that empirical objects are objects that we *can* cognize seems fundamental in this debate and essential to any reading of Kant’s view – it is shared in common between transcendental realists and transcendental idealists. So, again, it is very hard to see how this reading of TR could be accurate.

Setting this aside, let us move on to the next step: from the claim that we naturally regard empirical objects as positive noumena, together with the claim that ‘for Kant, noumena in the positive sense are members of an intelligible world’, Willaschek seems to derive the conclusion that ‘we must represent an intelligible world as a rational order, because we must project our own intelligence onto it and think of it as the creation of a divine (and thus supremely rational) intellect’ (p. 143).

This move goes by fast, and seems to smuggle something like the classical *logos* and *imago dei* doctrines into Kant’s vision of TR. Positive noumena are the things as God knows and creates them, the ‘divine mind is supremely rational’ and so ‘the world created and cognized by it must be a rational order’. Thus positive noumena constitute a rational order. If we mistakenly but naturally think that empirical objects are positive noumena, then we are mistakenly but naturally thinking of them as having this divinely ordained rational order. And then we leap from there to the assumption (via some sort of ‘projection of our own intelligence’) that we can cognize that order too. Willaschek thus provides another gloss on TR according to which it is the view

that empirical objects are part of a rational order and that human reason is therefore able to cognize these objects and their relations (even though this cognition will always be

imperfect, due to the imperfections of human reason compared with divine reason). (p. 144)

There is a lot to ask about in these last few steps, but one question is just how we went (naturally or not!) from thinking of positive noumena as constituting an intelligible, rational order for *divine* knowing to thinking of that order as one that 'human reason is therefore able to cognize'. Perhaps something like the ancient *logos* doctrine is at work here: the doctrine according to which all of reality is created by God with an intelligible, rational structure. But without the further *imago dei* doctrine (i.e. the assumption that our reason is structured analogously to the divine reason that structures the world and is thus able to track the order of that world), I do not see how one could get to the conclusion that 'human reason is therefore able to cognize' this order. And Willaschek does not explain the 'projection' step.

Kant himself rejects the *imago dei* doctrine: he thinks that God's way of knowing is categorically different from ours, and this leads him into all sorts of anxieties about how we can assume that the world is cognitively susceptible to understanding by minds like ours (this is a central theme of the third *Critique*). But is Willaschek suggesting that reason and common sense just naturally assume something like the *logos* and *imago dei* doctrines by way of this rather complicated set of inferences? Is the projection just something we naturally do? How would we adjudicate this?

Immediately following the passage just quoted, he offers another version of TR:

TR_{corresp}: 'There is a necessary correspondence between the principles of reason and the principles of reality' (p. 144).

Here Willaschek presumably means *human* reason. But note that there is an additional modal operator in the principle that was not part of the gloss leading up to it. The gloss spoke of 'imperfections', but in the principle itself the claim is that the correspondence between our rational way of thinking about reality and the way reality is *is necessary* rather than contingent – a strong claim indeed, and a far cry, it seems, from the idea back in A490/B518 that spatiotemporal features are mind-independent.¹

It is from TR_{corresp} that Willaschek goes on to argue that TR and transcendental illusion are 'two sides of the same coin' (p. 164). He has already shown that Kant thinks reason has a natural tendency to 'go beyond' and

take principles that ‘work fine’ in one domain (the empirical) and employ them in non-empirical domains. So now TR_{corresp} – which is also supposed to be something to which we are naturally and naïvely drawn (although I have just been raising questions about that) – is used to *justify* that tendency along with all sorts of ‘constitutive’ claims about non-empirical reality. In particular, it is used to justify the constitutive use of what Kant calls the ‘Supreme Principle of Pure Reason’:

When the conditioned is given, then the whole series of conditions subordinated one to the other, which is itself unconditioned, is also given (i.e. contained in the object and its connection).
(A307–8/B364)

Or, in Willaschek’s rendition,

Supreme Principle: ‘For all x, if x is conditioned, there exists the totality of conditions of x, which is itself unconditioned’ (p. 102).

So the bogus constitutive use of ‘the Supreme Principle (and with it the transcendental illusion at the heart of traditional metaphysics) rests on the tacit assumption of TR_{corresp}’ (p. 147).

By way of summary, here is the list of my concerns about how we arrived at this key version of TR – namely, TR_{corresp}:

1. The textual concern about where TR_{rationalism} comes from as a reading of TR, and how it relates to TR_{things}.
2. The minor question about whether we can assume that things in themselves are ‘intelligible objects’ or ‘noumena’ (p. 140).
3. The concern about ascribing a non-starter principle like TR_{pos} – interpreted as the claim that empirical objects are non-sensible and only cognized by God – to common human reason.
4. The concern about the move from thinking of empirical objects as noumena in a positive sense to thinking of them as part of a divinely instituted rational order.
5. The concern about the assumption that a divinely instituted rational order is one that ‘human reason is therefore able to cognize’ albeit in an ‘imperfect’ way (p. 144).
6. The concern about the sudden move from the claim that we are imperfectly able to cognize the divinely instituted rational order to the claim that there is a *necessary* correspondence between our rational principles and the principles of reality.

4. Why does the Supreme Principle Fail to Hold for Appearances?

Willaschek ends this key chapter by discussing a central question: 'Why does the Supreme Principle hold for things in themselves but not for appearances?' (p. 152). The problem is not in showing that the Supreme Principle holds for the things, but rather in showing that and why the Supreme Principle *does not* hold for the appearances. Here is the argument that it holds for the things:

P1: If x is R-conditioned, then there is at least one R-condition of x .

C1: If x is R-conditioned, then there is the totality of R-conditions of x .

P2: If y is the totality of R-conditions of x , then y is R-unconditioned.

C2: If x is R-conditioned, then there is the totality of R-conditions of x , which is unconditioned (= supreme principle).

P1 is supposed to be analytic, and the inference to C1 is a conceptual one that presupposes a naïve principle in set theory. In Willaschek's words, the inference goes through 'if we take Kant's definition of "totality" to express the naïve principle of set formation (sometimes called the principle of comprehension)' (p. 95). This is just the principle that if there is one thing that has F , then there is a non-empty set of all things that have F . Likewise if there is one R-condition of x , then there is a set of all the R-conditions of x – the totality of conditions.

P2 is also analytic: it is just the claim that the collection of all R-conditions of an x (all of x 's causes, let us say) is not itself R-conditioned. Thus the *totality* of causal conditions of this table in front of us is not itself caused. (Compare Kant: 'the totality of conditions is always itself unconditioned': A322/B379). And C2 (which is another formulation of the Supreme Principle) clearly follows from C1 and P2.

Willaschek says that we should not be surprised that one can prove that the Supreme Principle applies to things in a conceptual argument from analytic principles, given TR_{corresp}. He also says that, for appearances, it looks like P1 will not be analytic but will still be true given what Kant says about the principles of pure understanding in, for example, the Analogies. This is a small point, but I am not sure I understand it. Suppose R is a causal relation; then P1 says 'If x is causally conditioned, then there is at least one causal condition of x '. This still seems analytic

to me, contrary to what Willaschek suggests (p. 154). But maybe there are other sorts of conditions for which P_1 is more clearly synthetic. In any case, Willaschek agrees that P_1 is still true of appearances, and so if P_2 is analytic then it looks like we have a sound argument for the Supreme Principle as *applied to appearances*. Something has clearly gone wrong, since the whole point of Kant's critique of speculative metaphysics involves *rejecting* the claim that the Supreme Principle holds of appearances!

Willaschek himself suggests one way out: deny the set-theoretic 'principle of comprehension' that gave us the inference from P_1 to C_1 . But he admits that this would be anachronistic, and that in any case the restriction of the principle would 'have nothing to do with the distinction between appearances and things in themselves' and so we would have to deny it for things in themselves too, thus imperilling the whole picture (p. 155). At this point, he simply says that he will 'leave the question of how Kant can deny that the Supreme Principle holds for appearances unanswered' (p. 155). This is a very substantial thread that the book leaves untied.

I would like to conclude this part of my comments by offering a different way to tie it. If we look at Kant's own arguments against some of the specious metaphysical inferences, such as the cosmological one, we see him arguing that transcendental idealism is what allows us to block the argument:

On the contrary, if I am dealing with appearances, ... then I cannot say with the same meaning that if the conditioned is given, then all the conditions (as appearances) for it are also given; and hence I can by no means infer the absolute totality of the series of these conditions. For the **appearances**, in their apprehension, are themselves nothing other than an empirical synthesis (in space and time) and are thus given only **in this synthesis**. Now it does not follow at all that if the conditioned (in appearance) is given, then the synthesis constituting its empirical condition is thereby also given too and presupposed; on the contrary, this synthesis takes place for the first time in the regress, and never without it. (A498–9/B527)

A complicated passage, to be sure, but Kant at least *seems* to say that the transcendental idealist claim that empirical objects have their spatiotemporal features in a mind-dependent way allows us to resist *not* P_1 but rather the move from P_1 to C_1 . My own effort to analyse what it is for a feature to be merely phenomenal, for Kant, goes like this:

Mere Phenomenality: For any feature F that is not analytically mind-dependent,² any object O in world w , and the set of finite cognizers in w , ψ , O 's being F is *merely phenomenal* if and only if it is impossible that **both** (i) O is F and (ii) no member of ψ is *able to cognize* that O is F .

This principle (in my view) applies to all the spatiotemporal features of appearances, though *not* to their *existence* – they still exist in a mind-independent way.³

If this suggestion is in the right ballpark, it would allow us to resist the inference from P_I to C_I regarding appearances: for the *totality* of conditions will in many cases be indefinitely or even infinitely large, such that no actual finite mind can cognize it. And if no actual finite mind can cognize it, then it is not merely phenomenal. And if it is not merely phenomenal, then it does not exist in the domain of appearances. That may be a bit too anti-realist for Willaschek himself (given what I know about his earlier work), but it is not the ‘extreme form of idealism’ that Paul Guyer ascribes to Kant, according to which the *existence* of empirical objects is mind-dependent (p. 155).

There is more to say here, but even on this quick sketch I am curious whether Willaschek could accept this account of why the Supreme Principle does not apply to appearances. We have already seen that he thinks TI ‘minimally implies that all properties of objects we can cognize (“appearances”) depend (in some appropriate sense) on the possibility of being represented by finite rational beings like us’ (p. 247). That looks a lot like the analysis in *Mere Phenomenality* above. So even if he is not attracted to that kind of modest anti-realism himself, and even if he is not attracted to any form of TI himself, it seems like Willaschek could use something like this suggestion to block the inference from P_I to C_I for appearances on Kant’s behalf. An infinite totality of conditions cannot even in principle be cognized by an actual finite mind, and so we need not infer from a condition or a series of conditions to the empirical existence of a totality of conditions.

5. Denying Transcendental Realism without accepting Transcendental Idealism?

Finally, a few words about the very creative chapter 9, in which Willaschek argues that we can reject TR in all its forms without adopting TI (construed as the claim that empirical objects are not things in themselves).

This in turn allows him to ‘share Kant’s critique of speculative metaphysics without accepting his transcendental idealism’ (p. 244).

My first question is just one of clarification. Willaschek lays out four versions of TR and says that his goal is to show how we can reject them without accepting TI. Here are the first three, each of which we have encountered before:

TR: Empirical objects are things in themselves.

TR_{pos}: Empirical objects are noumena in the positive sense.

TR_{corresp}: ‘There is a necessary correspondence between the principles of reason and the principles of reality’ (p. 144).

But recall that TI is explicitly characterized as

TI: Empirical objects are not things in themselves.

Given this, it is hard to see how one can reject (i.e. negate) TR as articulated here without endorsing TI.

This just reveals, I think, that Willaschek’s real target in chapter 9 is not TR characterized in this most generic way. Rather, he is really targeting TR_{corresp} and another version which we have not looked at yet:

TR_{rep}: ‘Necessarily, if some object *o*, in order to be represented (by finite beings like us), must be represented as being *F*, then *o* is *F*’ (p. 244).

First, how do we get TR_{rep}? Willaschek says that it is a different thesis than TR_{corresp}, but still ‘closely related’ since TR_{corresp} follows from TR_{rep}. But recall that earlier the argument went from TR_{pos} to TR_{corresp}. So if TR_{rep} does not follow from either of these, it is not entirely clear what the argument for it is supposed to be. Ultimately, though, Willaschek wants to say that *this* is the ‘most general expression of the fundamental meta-ontological background assumption that Kant calls transcendental realism’ (p. 246).

Second, setting aside the conceptual basis or *origin* of TR_{rep}, note again here that we have got a necessity operator. We encountered that in TR_{corresp} above and noted that it was hard to see where it came from in the argument by which we derived that principle from TR_{pos}.

Third, Willaschek goes on to say that both TR_{rep} and $TR_{corresp}$ are in effect articulating a kind of 'epistemological realism': the thesis that 'our cognitive capacities are apt for providing us with knowledge of reality' (p. 246). But 'apt' has to be read very strongly to get this equivalence, since pretty much any non-sceptic is going to say that our cognitive capacities are 'apt' in the sense of *capable* of providing some knowledge of reality. In fact, it looks like the formulations of TR_{rep} and $TR_{corresp}$ (with the necessity operators in them) say that our faculties (or in the case of $TR_{corresp}$, our *rational* faculties) are not just apt but somehow *infallible*, at least when they actually succeed in representing an object. But now it just feels like we are setting up a straw man: does the metaphysician in a contemporary context need anything *that* strong? Even if we restrict it to our rational faculties (as $TR_{corresp}$ does), why would a metaphysician need to endorse the idea that *necessarily*, any time we succeed in representing something, we represent it accurately?⁴

We see the importance of the subtle addition of the necessity operator when we look at how Willaschek proposes to argue against TR (without endorsing TI). He first argues against TR_{rep} in its unrestricted form by appealing to optical illusions: Müller-Lyer and so forth. But I doubt that anyone is ever tempted by the unrestricted form of TR_{rep} , and in any case there was no argument for it (recall that it implied but was not implied by $TR_{corresp}$).

Willaschek goes on to talk about a restricted form of TR_{rep} where we just apply it to 'discursive judgments, beliefs, and thoughts' but says that even here it is 'hardly a plausible claim' because 'it's possible that, because of the limited capacities of our minds, we as finite cognizers necessarily form false beliefs about some aspects of nature' such as that bodies must either consist of indivisible parts or be infinitely divisible, when 'in fact – in a way we cannot fully comprehend – neither option is true' (p. 250).

Likewise he wants to reject $TR_{corresp}$ as a 'daring metaphysical thesis that faces many objections' (p. 251). Here one of his main arguments is that there is natural and moral evil in the world, and that this is 'contrary to reason' and thus challenges the idea that nature is a 'rational order'. But I think we would have to get much more specific about what it means for evil to be 'contrary to reason' in order for this argument to go through. Natural evil, like hurricanes and cancer, is surely able to be accounted for via natural causal explanation and so is not contrary to reason in that sense. Moral evil can be accounted for by free choices, and so while it

may be ‘irrational’ in some practical sense, its existence need not challenge something as general as TR_{corresp} .

Far more importantly, TR_{corresp} is only a ‘daring metaphysical thesis’ if it has the (unmotivated) necessity operator; otherwise it is a much less dramatic claim about aptness that many non-sceptics would find attractive. Willaschek suggests that, unless we follow Michael Della Rocca and generate an argument for the PSR, it is ‘difficult to think of a way in which [TR_{corresp}] could possibly be justified’ (p. 251). I agree that the version of the principle with the necessity operator is hard to support, but, again, why is *that* what is required by the contemporary metaphysician?

Remember the work that TR_{correp} was doing in the argument for the Supreme Principle. We have these natural tendencies to extend principles that ‘work fine’ in empirical contexts into the realm of the supersensible. TR_{corresp} – which again is supposed to be something to which we are naturally and naïvely drawn – is then wheeled in to *justify* that tendency and allow us to make all sorts of ‘constitutive’ claims about non-empirical reality. In particular, it justifies the bogus constitutive use of the supreme principle.

But if we go back to the *gloss* of the principle that Willaschek provides before stating TR_{corresp} , we find a much weaker claim. There (as we have already seen) he said that it is the view that ‘empirical objects are part of a rational order and that human reason is therefore able to cognize these objects and their relations (even though this cognition will always be imperfect, due to the imperfections of human reason compared with divine reason)’ (p. 144). But that is a *much* more modest principle than TR_{corresp} or TR_{rep} . Interpreted in this weaker way, I do not see how it could support the constitutive use of the Supreme Principle. And I also do not think it could be undermined by appeal to Müller-Lyer illusions or quantum mechanics or moral evil and the like.

In short, I think that the path that Willaschek takes from TR and TR_{pos} to $TR_{\text{correpond}}$ does not warrant the necessity operator that is crucial to making the latter (as well as TR_{rep}) seem like such an objectionable principle. If the principle just says that our rational faculties are apt to represent the structure of reality – albeit fallibly and imperfectly – then we not only do *not* need to endorse TI in order to reject it, we do not need to reject it at all.

Notes

- 1 This statement is also slightly different from TR_{rationalism} above. I am not sure whether anything important changes when we talk of principles rather than structure.
- 2 By this I mean any feature that is not dependent on a mind by its very concept – the way that ‘being thought about’ would be, for instance.
- 3 See my ‘Kant’s One-World Phenomenalism: How the Moral Features Appear’, in K. Schafer and N. Stang (eds), *The Sensible and Intelligible Worlds* (Oxford: Oxford University Press, 2020).
- 4 This is not to suggest, of course, that Descartes and Spinoza are not tempted to say something like this, at least about clear and distinct or accurate ideas. But in Chapter 9 Willaschek is operating for the most part in a contemporary context.