

Replies to the Comments of Paul Guyer and Andrew Chignell

MARCUS WILLASCHEK
Goethe-University Frankfurt am Main
Email: willaschek@em.uni-frankfurt.de

I would like to thank Andrew Chignell and Paul Guyer for their rich and challenging comments, which give me the opportunity to further explain and improve some of the theses and arguments of my book.¹ Before replying to their comments and for the convenience of the reader, I provide a brief précis of the book.

1. Précis of *Kant on the Sources of Metaphysics: The Dialectic of Pure Reason*

Kant's Rational Sources Account in the Transcendental Dialectic

In the *Critique of Pure Reason*, Kant asks: 'How is metaphysics as a natural predisposition possible? i.e. how do the questions that pure reason raises, and which it is driven by its own need to answer as well as it can, arise from the nature of universal human reason?' (B22). There are three theses implicit in Kant's questions:

RS-1 Rational reflection (about empirical questions) necessarily raises *metaphysical questions*.

RS-2 Rational reflection (by 'pure reason') on these metaphysical questions necessarily leads to *metaphysical answers* that appear to be rationally warranted.

RS-3 These metaphysical questions and answers arise from 'universal human reason', that is from rational thinking as such.

Following Kant's claim that the 'sources of metaphysics' lie in pure reason (cf. Axii, A309/B366), I call the conjunction of these three claims 'Kant's account of the rational sources of metaphysics' or the *Rational Sources Account*, for short. Kant extensively argues for the Rational Sources Account in the Transcendental Dialectic, which thus is not only a critique of speculative metaphysics about soul, world and God, but also a highly original reconstruction of that metaphysics out of the very structure of

rational thinking. Since these two sides of the Transcendental Dialectic are not always clearly separated by Kant, one central aim of my book (henceforth, *KSM*) is to reconstruct the Rational Sources Account and clearly distinguish it from the other, critical side of the Transcendental Dialectic.

The second central aim of the book is to show that the Rational Sources Account, while formulated in a philosophical framework that may seem foreign to many current philosophers, offers a highly attractive account of metaphysical thinking and is defensible even from a present-day perspective. In this latter respect, I argue that Kant highlights three features of reason that jointly explain the tendency to engage in speculative metaphysics: *discursivity*, *iteration* and *(the striving for) completeness*. Human reason is discursive, according to Kant, in that it does not take in truths ‘intuitively’ (holistically, at a glance), but by following stepwise procedures such as the progression from premises to a conclusion or, conversely, from something ‘conditioned’ to its ‘condition’. These procedures are *iterative*, in that they allow e.g. for pro- and episyllogisms and for asking for the *condition* of the condition of something conditioned, and so on. But at the same time, reason strives for *completeness* in that it is not satisfied until it has found supreme principles as highest premises and the totality of conditions of something conditioned. I argue that rational thinking indeed exhibits these features and that Kant is correct to claim that they naturally take us into metaphysical speculation.

Concerning the first aim, reconstructing the Rational Sources Account, Kant develops that account at *four levels* that roughly correspond to the four parts of the Transcendental Dialectic (Introduction, Books 1 and 2, Appendix):

- (1) the derivation of the ‘supreme principle of pure reason’ from a ‘logical maxim’ in the Introduction (A307–8/B364–5; *KSM*, chapters 2–5);
- (2) the derivation of the concept of the *unconditioned* from the logical form of rational inferences (A322–3/B378–9) and of the three classes of transcendental ideas (psychological, cosmological, theological) in Book 1 of the Transcendental Dialectic (A322–3/B379–80; A333–4/B390–1; *KSM*, chapter 6);
- (3) the dialectical inferences of pure reason (paralogisms, antinomies, arguments for God’s existence) and their conclusions (A348–642/B406–670), including the derivation of the transcendental ideas in Book 2 of the Transcendental Dialectic (*KSM*, chapters 7–8);

- (4) the transition from the logical principles of continuity, specification and affinity to the corresponding transcendental principles in the Appendix (A642–704/B670–732; *KSM*, chapters 4.2, 8.2).

Now at each of these four levels, we can discern the following threefold pattern, which gives us a general template of Kant's argument for the Rational Sources Account throughout the Transcendental Dialectic.

- (1) A *logical/real* (or *logical/transcendental* or *logical/metaphysical*) *transition*. It is rationally necessary to move from the 'logical' to the 'real' use of reason, where the 'logical' use only concerns cognitions as such and abstracts from content/objects, while the 'real' or 'transcendental' use concerns not just cognitions, but their objects. Since the logical use of reason aims at the 'unity of reason' (a system of cognitions), which requires *completeness* of cognitions (something we can never fully achieve), this means that at any given time we need *additional* cognitions (beyond the ones we already possess), which in turn requires the *real* use of transcendental ideas and principles (i.e. their use applied to objects).
- (2) A *regulative/constitutive transition*. It *appears* to be rationally necessary to move from the merely regulative to the constitutive use of reason (this appearance being what Kant calls 'transcendental illusion'). While the regulative use of reason's principles and ideas is merely heuristic and does not bring with it any metaphysical commitments, their constitutive use consists in taking them to be true representations of (the supersensible principles of) reality.
- (3) The transition from (or lack of distinction between) the regulative to the constitutive is motivated by the tacit assumption that the principles of reason necessarily correspond to the principles of reality ('transcendental realism').²

This threefold pattern explicitly or implicitly structures each of the four levels at which Kant argues for the Rational Sources Account, thus revealing a deep structural unity in the Transcendental Dialectic.

Some More Specific Claims

After an overview of Kant's conceptions of reason and metaphysics (chapter 1) and a detailed analysis of his account of a 'logical use' of reason (chapter 2), in chapter 3 I discuss a central element of Kant's account of speculative metaphysics, namely 'real' conditioning relations such as part/whole, inherence, causation, etc. As I understand Kant, something is conditioned in this sense if it is an object, property or event

that stands in need of explanation in some respect. Thus a composite object is explained (at least in one respect) by its parts, a property by its bearer and an event by its cause. So Kantian ‘real conditioning’ is similar to, but not quite the same as (what today is called) ‘metaphysical grounding’. However, I argue against Eric Watkins that real conditioning according to Kant is not a unified phenomenon, but falls into the three distinct classes (based on the relational categories) that Kant calls inherence (e.g. substance/attribute), dependence (e.g. causation, part/whole) and concurrence (e.g. interaction, ‘community’), and that while all these conditioning relations are explanatory and transitive, only some are asymmetrical and irreflexive. In the same chapter, I also show that there are three conceptions of the *unconditioned* at work in the Transcendental Dialectic. Put roughly, the first is that of an *unconditioned condition*, the second that of a *complete series of conditioned conditions* and the third, which is the most general one, that of a *totality* of conditions for something conditioned.

In chapter 4, I turn to what I call the ‘Transition Passage’ (A307/B364–5), where Kant claims that we must proceed from a ‘logical maxim’ (which Michelle Grier calls P₁) to the ‘supreme principle of pure reason’ (Grier’s P₂). While the ‘logical maxim’ concerns ‘logical’ conditioning relations (i.e. inferences) between cognitions, the ‘supreme principle’ concerns ‘real’ conditioning relations between objects, properties and events. Note that both principles exhibit the three features of reason mentioned above:

Logical Maxim (paraphrase): if a conditioned cognition is given, search for its condition (*discursivity*), and then for the condition of the condition (*iteration*), until you find some unconditioned cognition (*completeness*).

Supreme Principle: ‘when the conditioned is given, then so is the whole series of conditions [*discursivity*] subordinated one to the other [*iteration*], which is itself unconditioned [*completeness*], also given (i.e., contained in the object and its connection)’ (A307–8/B364).

Now Kant claims that ‘this logical maxim can become a principle of *pure reason* only through one’s assuming’ that, for everything conditioned, there is something unconditioned (A307–8/B365–6). Note that he does not say that we *mistake* the ‘logical maxim’ for the ‘supreme principle’ (which is how Michelle Grier and many others read him), but rather that the ‘supreme principle’ is the principle that the ‘logical maxim’

becomes by our assuming that, for everything conditioned, there is something unconditioned. Based on a close reading of some passages from the Appendix to the Transcendental Dialectic, I argue that Kant's use of 'assuming' here is deliberately ambiguous between a regulative sense (assuming something as a hypothesis) and a constitutive sense (assuming something to be true) and that he fully disambiguates between these two senses only 300 pages later in the Appendix. There, it finally becomes clear that using the 'supreme principle' and other transcendental principles constitutively (i.e. assuming them to be true) rests on a transcendental illusion and that only their regulative use (i.e. employing them as mere research hypotheses without committing ourselves to their being true) is legitimate.

In chapter 5, I discuss Kant's account of transcendental illusion. On the reading I develop, transcendental illusion consists in mistaking regulative principles and ideas for constitutive ones and rests on the tacit assumption of transcendental realism (TR), which Kant defines as the identification of, or lack of distinction between, appearances and things in themselves (A490/B518). Since the speculative metaphysician, whose thinking Kant reconstructs, thinks of 'things in themselves' as noumena in the positive sense, that is, as the objects of a divine and supremely rational perspective, TR becomes the claim that there is a necessary correspondence between the principles of reason and the principles that structure empirical reality (TR_C). Versions of this claim can plausibly be attributed to central figures in the history of Western metaphysics as different as Aristotle and Descartes. I argue that Kant thinks of transcendental illusion (mistaking regulative for constitutive principles) as motivated by the tacit assumption of TR_C.

In chapter 6, I claim (among other things) that there are exactly nine transcendental ideas (the soul as substance, as simple, as unity, as spiritual; the world as containing complete composition, division, origin, dependence; and God as *ens realissimum*), while the ideas of soul, world and God *simpliciter* are not transcendental ideas in Kant's technical sense. Relatedly, I locate the so-called 'metaphysical deduction' of the transcendental ideas (their derivation from the forms of the syllogism: categorical, hypothetical, disjunctive) not in Book 1 of the Transcendental Dialectic (On Transcendental Ideas), but in Book 2 (i.e. in the sections on the Paralogisms, Antinomies and Ideal), and I suggest that Book 1 mainly derives the concept of the unconditioned and the three 'classes' of transcendental ideas (psychological, cosmological, theological).

In chapters 7 and 8, I turn to Book 2 of the Transcendental Dialectic. With respect to the Paralogisms, I argue (among other things) that transcendental illusion in this case rests on what can be understood as a generalization of TR_C , namely TR_{rep} , which says: If, to be represented at all (by finite beings like us), some object o must be represented as being F , then o is F . For instance, in the thought ‘I think’, we must represent ourselves as substances, as simple, as person and as a spirit; thus, by TR_{rep} , we conclude (in the four paralogisms) that we are substances, simple, persons and spirits. Kant’s rejection of the paralogisms is motivated in part by his rejection of (the unrestricted validity of) TR_{rep} .

Concerning the Antinomy chapter, I show how all four antinomies rest on an inference that starts with a constitutive use of the supreme principle (*If there is something R-conditioned, there must also be something R-unconditioned*) and concludes that, since there is something that is conditioned in some respect (e.g. temporally, mereologically, causally, etc.), there must exist something that is unconditioned in that respect. Now the thesis-sides of the antinomies argue, each for a specific conditioning relation, that this unconditioned cannot be an infinite series of conditions and hence must be an unconditioned condition, while the antithesis-sides argue that the unconditioned cannot be an unconditioned condition and hence must be an infinite series. In this way, both sides of the antinomies are committed to the existence of something unconditioned. This commitment, however, is illusory since it rests on a constitutive use of the ‘supreme principle’.

Finally, concerning the Ideal of Reason, I reconstruct Kant’s derivation of the transcendental ideal (the concept of an *ens realissimum*), which, in keeping with the general pattern of Kant’s argument for the Rational Sources Account, *starts* with a transition from the ‘logical’ principle of determinability to the ‘transcendental’ principle of complete determination and *ends* with the claim that the latter is valid for experience (when used *regulatively*). Note that the three arguments for God’s existence Kant famously criticizes (ontological, cosmological, physico-theological) are not part of the Rational Sources Account, since Kant rejects them as ‘unnatural’, ‘an artifice’ and ‘dialectical’. The only ‘natural’ argument for God’s existence, which is part of the Rational Sources Account (A586–7/B614–15), is an abductive argument that first derives the conclusion that ‘there is at least one necessary being outside the world’ and then adds that the best possible candidate for a necessary being outside the world is the *ens realissimum*.

The final chapter (chapter 9) argues (*inter alia*) that one can reject TR without accepting Kant's own transcendental idealism. Thus Kant gives us a plausible genealogy of metaphysical thinking from its sources in human reason, but we do not have to become transcendental idealists in order to resist this kind of metaphysics.

The book closes with a brief postscript on Kant's 'practical metaphysics', where the moral law and the idea of a highest good serve as the basis for metaphysical claims (the 'postulates') about immortality, freedom and God.

2. Reply to Paul Guyer

Paul Guyer takes issue with three aspects of my reconstruction of Kant's account of the rational sources of metaphysics. First, he finds 'the framework of the book slightly confusing' with respect to the way I distinguish between Kant's critique of speculative metaphysics and what I call the Rational Sources Account. Second, he points out two respects in which the concept of God plays a more important role for Kant's account of metaphysics than I ascribe to it. And third, Guyer objects to my account of the highest good and the postulates based on it (in my brief postscript on Kant's 'practical metaphysics'). I agree with Guyer that none of these points amounts to a substantial disagreement. Still, I am grateful for the opportunity to clarify and further explain some aspects of my reading of Kant.

First, the general framework of my book: I distinguish *two strands* in the Transcendental Dialectic: (i) the critique of traditional speculative metaphysics and of its alleged proofs and (ii) the Rational Sources Account, which reconstructs this form of metaphysics as arising from the structure of rational thinking (discursivity, iteration, completion). The Rational Sources Account, as I understand it, consists of *three sub-theses*, and it is developed by Kant on *four levels* (which correspond to the four parts of the Transcendental Dialectic) in a way that follows a *three-part pattern* (a logical/real transition and a regulative/constitutive transition plus the assumption of transcendental realism; cf. the précis of the book above). I admit that this complex structure can be confusing. But I think that it is necessary to bring out the complexity of Kant's argument in the Transcendental Dialectic.

Guyer mainly complains that it is difficult to keep the two strands of the Transcendental Dialectic apart, the critical and the reconstructive. But this may be due to the fact that the critical strand has long been so

dominant that we automatically read some steps of the other strand, the Rational Sources Account, as critical. In particular, it is tempting to understand the tendency to use the principles and ideas of pure reason *constitutively* (and not just regulatively) as part of Kant's *critique* of metaphysics. Thus, Guyer writes: 'the Rational Sources Account is Kant's derivation of the regulative principles of reason from its most basic operations, while the critical core of the Dialectic is its explanation that we transform the legitimate regulative principles of reason into illegitimate constitutive principles because of our natural but unwarranted attraction to transcendental realism' (p. 6). But note that Guyer's 'critical core of the Dialectic' consists of two logically distinct points: (i) the transition from the regulative to the constitutive use of reason and (ii) the claim that the constitutive use is illegitimate. While the second point is part of Kant's critique of metaphysics, I think of the first point as part of the Rational Sources Account, since it is the regulative/constitutive transition that explains why we not only ask metaphysical questions and entertain metaphysical hypotheses, but tend to accept metaphysical theses and theories as true representations of reality. That this step is *illegitimate* is an additional and logically independent point Kant makes.

This means that we must distinguish two questions: (1) Why do we tend to use rational principles constitutively (Kant's answer: because of a tacit assumption of transcendental realism) and (2) Are we justified in using rational principles constitutively? (Kant's answer: no, because that would take us beyond the limits of possible experience). As I understand it, while the answer to the first question is part of the Rational Sources Account, the answer to the second is part of Kant's critique of metaphysics. Note that the Rational Sources Account is meant to be *compatible* with traditional metaphysics, since it reconstructs the thinking that underlies it and shows that it does not arise from contingent features of human psychology or from individual mistakes, but rather from the structure of reason as such.³

In sum, while Guyer is correct to say that the critical core of the Transcendental Dialectic lies in *criticizing* the transition from the regulative to the constitutive, this does not mean that this transition itself cannot be part of the Rational Sources Account. Once this is granted, I think it is possible to keep apart the critical and the reconstructive strands in the Transcendental Dialectic.

Let me add a remark concerning a more specific point that Guyer makes in the first part of his comments. Guyer reads the 'logical maxim'

(to *seek* the unconditioned for everything conditioned, A307/B364) as the regulative version of the ‘supreme principle’, which claims that there *exists* something unconditioned for everything conditioned. He implies that the ‘logical maxim’ therefore is not a logical principle ‘properly speaking’ (p. 3), since this would require that it abstract from objects and only concern cognitions. As I argue, however, the logical maxim is better understood as concerning not objects but conditioned *cognitions*, and thus requiring us to seek unconditioned *cognitions* (i.e. first principles), which is not only what Kant’s text suggests, but also allows for the ‘logical maxim’ to be a logical principle in Kant’s sense.

I now turn to Guyer’s ‘substantive suggestions’ (pp. 6ff). First, Guyer objects that God plays a more important role in Kant’s argument than I allow for: ‘The idea of God is thus not just the idea of one more unconditioned alongside others, but in Kant’s mind all roads lead to the idea of God’ (pp. 7–8). While I agree with this claim, I do not agree with Guyer’s reading of the passage on which he bases it. Kant writes that there are ‘three classes’ of transcendental ideas, ‘of which the first contains the absolute (unconditioned) unity of the thinking subject, the second the absolute unity of the series of conditions of appearance, the third the absolute unity of the condition of all objects of thought in general’ (A334/B391). Guyer glosses this as follows: ‘In other words, however the ideas of the unconditioned subject and world-whole might be reached, the idea of God is inevitably reached as the unconditioned condition of those’ (p. 7). If Guyer wants to say here that God is thought of as the unconditioned condition specifically of our substantial souls and the empirical world as a whole, I do not think that this is what Kant says in the quoted passage. The three kinds of the unconditioned are distinguished here according to the objects they condition, namely a subject’s representations in the first case, empirical phenomena in the second and objects of thought in general in the third. Of course, our souls and the world are also objects of thought and, according to the metaphysics Kant reconstructs here, conditioned by God. But it is not as if the idea of God was introduced specifically as the unconditioned condition of the objects of the other two ideas.

Next, let me briefly comment on a remark Guyer makes in passing: ‘Thus transcendental realism is blocked with regard to the first two antinomies, but allowed as an indemonstrable possibility for the theses of the last two’ (p. 7). I assume that Guyer here thinks of transcendental realism as the claim that there are things in themselves that ground empirical phenomena, since it is this claim that allows Kant to resolve the

dynamical antinomies. But that is not what Kant himself calls ‘transcendental realism’ and what I mean by that term in my book (cf. Chignell’s comments and my reply for a discussion of transcendental realism). As I see it, transcendental realism (TR) consists in the identification of empirical objects with things in themselves (and, more specifically, with noumena in the positive sense). This identification, which Kant of course rejects, is a background condition of all four antinomies, since they all rest on a constitutive reading of the ‘supreme principle’, which in turn is motivated by TR. Given the constitutive reading of the ‘supreme principle’, it follows from the existence of conditioned empirical objects that something unconditioned exists. The antinomies then arise because this ‘unconditioned’ can be thought of either as a first unconditioned condition or as the totality of conditioned conditions (A417/B445). Since the *resolution* of the antinomies is not part of the Rational Sources Account (but rather part of the critical strand of the Transcendental Dialectic), I do not discuss it in detail in my book. But here is how I think my reading would play out when applied to the resolution of the antinomies: If we give up TR, we lose the implicit premise of the theses and antitheses, namely that there is something unconditioned. In the first two antinomies, which concern only relations within the empirical world, this means that both sides turn out to be false (since series of conditions in the empirical world are neither finite nor infinite); in the last two antinomies, which potentially concern relations between events in the empirical world and their intelligible grounds as well, this means that both sides can be true, since there may be unconditioned conditions of empirical objects and events that lie outside the empirical world.⁴

Next, Guyer suggests that in the published introduction to the third *Critique* (5: 180), Kant introduces the idea of God not as resting on a form of transcendental realism (TR), as he does in the first *Critique*, but on a misapplication of transcendental idealism:

Kant is arguing that we naturally arrive at a conception of God by an extension of the fundamental insight of transcendental idealism with regard to the most general laws of nature, that they are the product of our own understanding, to all laws of nature, which, as laws, must be necessary and, as necessary, must be the product of an understanding, although one that is not ours and is obviously greater in capacity than ours. (p. 8)

I agree with Guyer’s reading of the passages from the third *Critique* and would only like to point out that, on my understanding of TR,

TR and Kant's own transcendental idealism are not contradictory, but merely contrary theses. The most general form of TR I discuss in my book is the claim (put roughly) that things are as we necessarily represent them to be (TR_{rep}). In the hands of the traditional metaphysician, this becomes the thought that the world must be as reason makes us expect it to be, a thought Kant himself rejects. But note that Kant's own transcendental idealism can be seen as a restriction of TR_{rep} to empirical objects: While it is false to assume that the world in general must be the way we necessarily represent it, according to transcendental idealism it is *true* that the empirical world corresponds to the necessary features of our representing objects (space, time, categories). But then, the thought from the third *Critique* that Guyer highlights (all laws as necessary come from some understanding that imposes them on its objects) is not an extension, or misapplication, of transcendental idealism, but rather a form of TR, according to which the necessary features of representing objects correspond to the necessary features (laws) of reality.

Finally, Guyer turns to my brief postscript on Kant's *practical* metaphysics, including the postulates of pure practical reason. I claim that because the postulates are theoretically undecidable, theoretical reason on its own would have to reject them; but because practical reason makes their acceptance necessary, theoretical reason must 'give in'. First, Guyer objects that theoretical reason does not have to 'give in', because undecidability is not rejection. I take him to mean that theoretical reason does not have to reject the postulates, since given their undecidability it can remain neutral. Thus it does not have to 'give in' to practical reason, because it is not *opposed* to the postulates in the first place.

But I think that this underestimates Kant's commitment to the traditional evidentialist view, found for instance in the ancient sceptics, Descartes and Hume, that if some proposition is theoretically undecidable, because we do not have conclusive evidence for or against it, the only rational response is suspension of judgement. Thus, considered from a purely theoretical point of view, we cannot rationally accept the postulates (5: 120). In this sense, theoretical reason must give in to practical reason: Since accepting the postulates is rationally necessary, theoretical reason must curb its inherent evidentialist tendency by allowing that we *accept* the postulates even though, on purely theoretical grounds, they would *not* be rationally acceptable.

Second, Guyer suggests that Kant's view of the highest good in the first *Critique*, according to which it can be realized only in an afterlife,

in later works gives way to a more plausible view, according to which the highest good is the *telos* of historical progress and can be realized, or at least approximated, in this world. I agree that we can find both conceptions in Kant's work and that the latter is by far the philosophically more attractive.

3. Reply to Andrew Chignell

Andrew Chignell focuses his comments on the formulations of transcendental realism (TR) I introduce in chapter 5 and on my criticism of TR in chapter 9. In sections 2 and 3 of his comments, he asks how the different formulations of TR I work with are related and how they can be derived from Kant's own formulation of TR as the identification of empirical objects with things in themselves (which he calls 'TR_{things}'). I will first offer a general response to Chignell's concerns and then answer two more specific questions he raises in this connection. Finally, I will briefly address the points he raises in sections 4 (concerning the supreme principle and appearances) and 5 (concerning my critique of TR).

First, on my reading of the Transcendental Dialectic, transcendental realism explains the phenomenon Kant calls 'transcendental illusion', which at the most general level consists in mistaking subjective principles and ideas for objective ones (A296–7/B353–4). I base this reading on Kant's claim that transcendental idealism is the 'key' to the resolution of the antinomies (A490/B518), which implies that transcendental realism explains the transcendental illusion involved in the antinomies. Since Kant does not offer any other explanation for the transcendental illusion involved in the paralogisms and the theological arguments, but seems to assume a unified account of transcendental illusion, it is plausible to start from the 'hypothesis' (KSM, p. 140) that transcendental realism explains transcendental illusion throughout the Transcendental Dialectic.

In order to make this hypothesis work, I ask what exactly Kant means by 'transcendental realism', which he officially introduces at A490/B519 as the identification of, or lack of distinction between, appearances and things in themselves (see also A369). A satisfactory interpretation of what Kant means by transcendental realism (TR) in the context of an explanation of transcendental illusion must satisfy three conditions: (C1) It must explain how the traditional metaphysician, by identifying empirical objects with things in themselves, is committed to TR in the relevant sense. (C2) It must explain how the tacit assumption of TR in that sense can plausibly serve as an explanation for transcendental illusion.

(C₃) It must be such that TR can plausibly be attributed (as a tacit commitment) to ‘universal human reason’.

In *KSM*, I argue that the best candidate for playing this threefold role is TR_C, which says that there is a necessary correspondence between the principles of reason and the principles that structure reality. The idea is that, given a tacit assumption of TR_C, one will naturally assume that the subjective principles of human reason (which, according to Kant, can be properly used only regulatively) hold for reality itself (and thus can be used constitutively).

Chignell primarily questions whether TR_C satisfies C₁, that is, whether TR_C is suitably related to Kant’s official definition of TR, and he raises doubts about several of the steps by which I arrive at TR_C. He assumes that my aim is to show that TR_C follows from, or can be derived from, Kant’s generic definition of TR as identification of appearances and things in themselves, and he questions the cogency of that derivation. I now see that some of my formulations indeed invite such a reading and I am grateful to Chignell for forcing me to clarify this point. In fact, my aim is not to derive one or more formulations of TR, but merely to explain how Kant could have thought that the identification of empirical objects and things in themselves could commit one to TR_C. Note that Kant does not have to claim that his generic formulation of TR (empirical objects are things in themselves) and TR_C are equivalent; it would be sufficient for my purposes if he held that the two formulations would appear to be equivalent to the traditional metaphysician (or, even weaker, that the traditional metaphysician is committed to TR_C insofar as she does not distinguish between empirical objects and things in themselves).

Now why these two claims should appear to be equivalent requires a complicated explanation (cf. *KSM*, pp. 140–8), but the main idea is that, in our context, things in themselves are thought of as noumena in the positive sense, which are the objects of a divine perspective, which in turn are the elements of a rational order, thus resulting in the claim that empirical objects are elements of a rational order which necessarily corresponds to the principles of reason. I do not claim that this is a cogent derivation of TR_C from Kant’s generic formulation of TR or even that TR_C is a plausible reading of what Kant meant at A490/B519 by ‘realism in the transcendental signification’, but merely that it was plausible for Kant to assume that in the hands of the traditional metaphysician, and given a set of background assumptions she is likely to accept, TR would amount to something like TR_C.⁵ Thus, condition C₁ is satisfied.

Moreover, I argue that only TR_C and not TR in its generic form (identification of empirical objects with things in themselves) can explain transcendental illusion (C_2). So I take it that TR_C satisfies at least two of the three criteria. Concerning the final condition – TR as part of common sense or universal human reason – I think that it is the most problematic. What needs to be shown here is that there is a tacit commitment to TR_C rooted in human reason (*KSM*, 148–50). To see that this is not at all implausible, consider that reason is a *cognitive* faculty, aimed at cognizing reality. Therefore, if reason makes us expect reality to be a specific way, for instance to exhibit a systematic order, we tend to *assume* that reality really is that way. Of course, according to Kant that assumption is unwarranted, but it appears natural enough. If, as I would like to claim, a tacit commitment to TR_C is our best explanation for this natural tendency, then Kant was right to attribute TR_C to universal human reason.

Thus my general response to the questions Chignell is raising in sections 2 and 3 of his comments is this: it is not my aim to derive TR_C from Kant's generic formulation of TR, but rather to argue that TR_C is the best candidate for a unified explanation of transcendental illusion throughout the Transcendental Dialectic. I do not think that his concerns undermine that aim.

Let me now turn to two of Chignell's more specific questions concerning TR. First, he asks whether TR_{pos} (empirical objects are noumena in the positive sense) is not an obvious non-starter, given that it identifies *empirical* objects with noumena in the positive sense, which are *non-sensible* objects. I do not think that this is the case, because Kant's conception of noumena in the positive sense allows for these objects also to have *sensible* properties, as becomes clear from the fact that human beings, considered as noumena in the positive sense, have an intelligible character, but *also* have an empirical character and other sensible properties (cf. A539/B567). In this way, empirical objects could be identical to noumena in the positive sense, which can be *fully* cognized only from a divine perspective. That is just what TR_{pos} claims.

Moreover, Chignell asks where the necessity operator in TR_C comes from. TR_C says that there is a *necessary* correspondence between the principles of human reason and the principles that structure reality. Let me first point out that this does not mean that, according to TR_C , our rational capacities are infallible, as Chignell suggests. After all, the correspondence in question may be necessary, but only partial; moreover,

we can make mistakes in following the principles of our own reason (as we do in the case of logical fallacies and, according to Kant, in the case of transcendental illusion), even if these principles should necessarily correspond to reality.

Now one way for a traditional metaphysician to arrive at TR_C (including necessity) is by assuming that God created the world as a rational order and created human reason as a means for us to discover that order. This assumption is the foundational idea behind medieval scholasticism and it was still common in Kant's own time. An alternative route to TR_C (including necessity) would be the Aristotelian claim that the forms of objects and the forms of true thoughts are identical. Again, it is not my claim that TR_C follows logically from Kant's generic definition of TR, but merely to explain why the fact that the traditional metaphysician does not distinguish empirical objects from things in themselves could plausibly result in the assumption that the principles of human reason necessarily correspond to reality.

In section 4 of his comments, Chignell offers an answer to a question I raise but do not answer (KSM, pp. 152–6) about why, according to Kant, the 'supreme principle of pure reason' (roughly: *if there is something R-conditioned, then there is the totality of its R-conditions, which is R-unconditioned*) does not hold for appearances, which is a claim Kant is committed to in the context of his resolution of the antinomies. Chignell's solution rests on the idea that, while appearances are 'merely phenomenal' in that all their properties qua appearances are mind-dependent, they nonetheless do not, according to Chignell's definition of 'mere phenomenality', depend on our minds for their existence. This allows for the following solution:

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 [according to the proffered definition] merely phenomenal. And if it is not merely phenomenal, then it does not exist in the domain of appearances. (p. 11)

But note, first, that this does not apply to the case of an unconditioned condition (since in that case the totality of conditions is not infinite); and second, even in the case of an infinite totality of conditions, it would only follow that the unconditioned does not exist in the domain of appearances, and not that it does not exist at all. But I take it that what

Kant needs to reject is not just a version of the ‘supreme principle’ that allows us to infer, from the existence of something conditioned in appearance, the existence of something unconditioned *in the domain of appearances*. Rather, he must reject a version that allows us to infer the existence of *something unconditioned* (without further qualification), since it is that (stronger) version on which the antinomies rely. I do not see how Chignell’s conception of the ‘merely phenomenal’ character of appearances helps us understand how Kant can reject that stronger principle.

In closing, Chignell addresses my claim that transcendental realism (in the form of TR_C) can be rejected without accepting Kant’s transcendental idealism. He objects that my argument is successful only if TR_C contains the (as he sees it) unmotivated necessity operator. He thinks that transcendental realism is in fact *true* if it is understood as the claim that our rational faculties allow us to cognize reality, albeit fallibly and incompletely. I agree that such a view is unobjectionable. But note that this view does not give rise to transcendental illusion (as Chignell is of course well aware), so it is very different from the kind of view that underlies traditional metaphysics and is criticized by Kant. What matters for the traditional metaphysician, as Kant understands him, is that our rational faculties *alone*, independently from empirical input and thus completely *a priori*, allow us to cognize the fundamental structures of reality. As I understand Kant, he thinks that this commits the traditional metaphysician to something like TR_C. It is TR_C, not Chignell’s more moderate version of transcendental realism, of which I argue that it can be rejected without accepting Kant’s own transcendental idealism.

Let me close by repeating that I am truly grateful to Andrew Chignell and Paul Guyer for their comments, which contain more important questions and suggestions than my brief response can do justice to.

Notes

- 1 The contributions to this book symposium were first presented at an Author Meets Critics session of the North American Kant Society at the Pacific APA Meeting in Vancouver, 19 April 2019. I am grateful to Eric Watkins for making this event possible and to Lucy Allais for chairing the session in Vancouver. I would also like to thank the editors of *Kantian Review* for their willingness to publish our exchange in this journal. Further thanks to Fabian Burt for valuable comments and to Sean Neagle to helpful linguistic advice.
- 2 See Andrew Chignell’s contribution to this symposium and my response for a discussion of the role of transcendental realism in Kant’s Rational Sources Account.

- 3 Relatedly, Guyer takes the three steps of the general template of Kant's Rational Sources Account that I reconstruct in chapter 4 of my book (a logical/real transition, a regulative/constitutive transition plus the assumption of transcendental realism) to represent the 'purely critical strand of Kant's argument' (p. 3). But as I see it, these three steps as such are not *critical* of metaphysics at all, which can be seen from the fact that a proponent of metaphysical realism might accept them as a fair statement of her view.
- 4 This presupposes that the antithesis is interpreted as not implying the existence of a totality of (an infinite series of) conditions, but only as denying the existence of an unconditioned condition within the empirical world.
- 5 Note, by the way, that TR_C is not meant to *justify* our tendency to use regulative principles constitutively, as Chignell (pp. 8, 14) assumes, but merely to *explain* it. Of course, Kant does not think that this tendency *can* be justified.