



This is a repository copy of *Explaining synthetic a priori knowledge : the Achilles' heel of transcendental idealism?*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/184365/>

Version: Accepted Version

---

**Article:**

Stern, R. [orcid.org/0000-0003-2967-647X](https://orcid.org/0000-0003-2967-647X) (2022) Explaining synthetic a priori knowledge : the Achilles' heel of transcendental idealism? *Kantian Review*, 27 (3). pp. 385-404. ISSN 1369-4154

<https://doi.org/10.1017/S1369415422000115>

---

This article has been published in a revised form in *Kantian Review* <http://doi.org/10.1017/S1369415422000115>. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © The Author(s), 2022. Published by Cambridge University Press on behalf of *Kantian Review*.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Explaining Synthetic a priori Knowledge: The Achilles' Heel of Transcendental Idealism?

Robert Stern

University of Sheffield, UK  
Email: r.stern@sheffield.ac.uk

## **Abstract**

This article considers an apparent 'Achilles' heel' for Kant's transcendental idealism, concerning his account of how synthetic *a priori* knowledge is possible. The problem is that while Kant's distinctive attempt to explain synthetic *a priori* knowledge lies at the heart of his transcendental idealism, this explanation appears to face a dilemma: either the explanation generates a problematic regress, or the explanation it offers gives us no reason to favour transcendental idealism over transcendental realism. In the article, I consider G. E. Moore's version of the problem, which I argue has not yet received an adequate response. Instead, I offer a way out of this dilemma by focusing on the normativity rather than the metaphysics of the mind (sections 3–6).

**Key words:** Immanuel Kant; G. E. Moore; synthetic *a priori* knowledge; transcendental idealism; transcendental realism; the Copernican revolution; epistemic norms

My aim in this article is to consider a worry that might be seen as an 'Achilles' heel' for Kant's transcendental idealism, concerning his account of how synthetic *a priori* knowledge is possible. This worry is not original to me, having been raised on several occasions by Kant's critics; but while various related issues have been widely discussed, the worry as such has not elicited much direct response from Kant's defenders.<sup>1</sup> This is curious, as the issue does not seem a trivial one, but rather could reveal a vulnerability that will allow Kant's critics to fatally wound his transcendental idealism, just as the weakness in Achilles's heel was ultimately to lead to his death. I think the worry is therefore serious enough to deserve an

explicit response: I will examine some options in what follows, including one that I will recommend as more promising than the rest.

The problem, put simply, is that while Kant's distinctive attempt to explain synthetic *a priori* knowledge lies at the heart of his transcendental idealism, this explanation appears to face a dilemma which it cannot escape: either the explanation generates a problematic regress, or the explanation it offers gives us no reason to favour transcendental idealism over transcendental realism. I will first say something about how this problem relates to Kant's overall project (section 1), before going into the problem in more detail (section 2). I will then consider some possible Kantian responses, one of which I think holds out some hope as a way out of this dilemma (sections 3–6).

### **1, Kant's project**

While of course Kant's project is multifaceted and many-sided, with a vast range, it can be argued that at the heart of it lies the puzzle of synthetic *a priori* knowledge, and Kant's proposed explanation of how such knowledge is possible. This centrality is clear from the priority Kant himself gives to this issue, both in the first *Critique*,<sup>2</sup> and also in the *Prolegomena*, where it is presented as the key question that requires resolution;<sup>3</sup> and it is correspondingly important to the success of that critical project, as Kant makes solving this issue a unique achievement of his transcendental idealism.

So how does Kant take this solution to work? The answer, of course, is his 'Copernican revolution': once we take the Copernican turn, and 'assume that objects must conform to our cognition' (*CPR* Bxvi) rather than the other way round, then we can understand how we can know *a priori* that e.g. 'everything which happens must have a cause', even though this is a synthetic proposition not an analytic one.

Why? Because we know that causality is part of our cognitive framework, in the sense that in so far as we experience an object, it must comply with the 'rules' of that framework, and thus must conform to this principle of causality. Of course, it is possible that outside our

experience, there are objects that do not obey this principle – but then, we could never experience them, and they are never going to falsify the principle as far as our experience is concerned. We thus get the universality and necessity of our claim within the world as far as we could ever experience it, which is enough to give us synthetic *a priori* knowledge of this world. And while some may wish for more, and so feel disappointed with this result, these transcendental realists are hankering for what is beyond us as finite knowers, which is knowledge of ‘things in themselves’ outside appearances – whereas what Kant gets us is cognition of things as they appear to us, about which we can have securely grounded synthetic *a priori* knowledge, based on the Copernican principle that ‘we can cognize of things *a priori* only what we ourselves have put into them’ (*CPR* Bxviii). At the same time, Kant’s position does not have to claim that *everything* in the world we experience is mind dependent, but only the ‘formal’ structures of space and time and the categories, so the ‘material’ idealism of a Berkeley is thus avoided. The Copernican turn thus appears to brilliantly explain how such synthetic *a priori* knowledge is possible, while avoiding the pitfalls of either empirical idealism or transcendental realism, which would leave such knowledge mysterious.

Now, it is difficult not to be charmed by the elegance of this solution to the problem of synthetic *a priori* knowledge, and many have been. But critics claim to have found its Achilles’ heel.

## **2. The problem**

As I have said, I am not the first to identify this problem. In fact, it might be lurking in the background of the various post-Kantian worries about Kant found in the work of the later German idealists,<sup>4</sup> and other subsequent critics. But the clearest form in which I have found the problem expressed is in this passage from G. E. Moore’s *Some Main Problems of Philosophy*, and that is how I will introduce it here.<sup>5</sup>

Moore first sets out Kant's solution to the problem of synthetic *a priori* knowledge, using the example of the angles of a triangle:

What we are here concerned with is the answer which he gives in the case of one particular class of universal synthetic propositions. Namely propositions like this:

Three angles of a triangle are always equal to two right angles ... [H]is answer to this question ... is as follows.

He says that we could not possibly know them, if Space were anything else than a mere form in which, owing to the constitution of our minds, things appear to us. If our minds are so constituted that the angles of every triangle which appears to us are always equal to two right angles, then we may be quite certain that every triangle which ever does appear to us will have the sum of its angles equal to two right angles.

Moore then goes on:

This is, of course, true. But does this, in fact, explain at all, how we can *know* that this always has been and always will be the case? Obviously, it does not, unless it first be explained or be self-evident, how we can know that our minds *are* so constituted as always to give this result. But this proposition, that our minds are so constructed as always to produce the same appearances, is itself a universal synthetic proposition, of precisely the sort with regard to which Hume pointed out the difficulty in seeing how we can know them. When Kant assumes that we do know that our minds *are* so constituted, he is assuming that each of us does know not only that his own mind always has acted in a certain way, but that it always will do so; and not only this but also that the minds of all other men always have acted and always will act in this way. But how can any of us know this? Obviously it is a question which requires an answer just as much as any of those which Kant set out to answer, and yet he never even

attempts to answer it: it never seems to have occurred to him to ask how we can know that *all* men's minds are so constituted as *always* to act in a certain way. And once this question is raised, I think the whole plausibility of his argument disappears...

It is not, in fact, a bit plainer how you can know universal synthetic propositions about the action of the human mind, than how you could know them about other things; and hence the argument that anything about which you can know universal synthetic propositions must be due to the action of the human mind, entirely loses all plausibility. (Moore 1953: 152–4)

I take it Moore's objection can be put as follows:

1. Kant explains how synthetic *a priori* knowledge is possible by appealing to the necessary structures of our minds
2. But this claim about the necessary structure of our minds is itself a form of synthetic *a priori* knowledge
3. But Kant does not face up to the difficulty of explaining how this knowledge of our minds is possible
4. Nor can this be explained on the Kantian approach
5. So Kantian idealism fails to explain how synthetic *a priori* knowledge is possible

Let us consider the argument in a little more detail:

Premise 1 seems to be a reasonable way to take Kant's Copernican revolution, as we have discussed.

Premise 2 also seems plausible: As a necessary claim about how our minds work, it would on Kantian grounds appear to be *a priori* – for as Kant says 'if a proposition is thought along with its *necessity*, it is an *a priori* judgement' (*CPR* B3). And it is hard to see how it could be analytic – though this is a possible response to Moore's objection which I will return to in what follows.

Premise 3 is a textual claim, pointing to the fact that Kant never seems to pose for himself the challenge that Moore raises, and try to deal with it explicitly, by asking how his Copernican revolution might apply at a second-order level and hence explain how we can have synthetic *a priori* knowledge of how the human mind is necessarily constituted. As far as it goes, I think this claim is correct, in that Kant does not spell out Moorean concerns and then attempt to deal with them directly. However, this in itself may not be a problem, as Kant may have reasons to think such concerns just do not arise rather than have overlooked them, so that this textual point may not tell us much taken on its own.

Premise 4 is a more systematic claim: that Kant's approach makes it 'no easier' to see how this second-order synthetic *a priori* knowledge of the necessary structure of our minds is possible than is first-order synthetic *a priori* knowledge of the world outside our minds. Moore does not really spell out his concern here, but I take it the issue is as follows:

Kant attempts to explain our first-order synthetic *a priori* knowledge by appeal to the necessary structures of our minds; but this just involves a second-order synthetic *a priori* knowledge concerning that structure. But how is that knowledge to be explained? It cannot itself be explained by a further appeal to the structure of our minds, and so it is left just as unexplained as the first order synthetic *a priori* knowledge was initially, so we have got no further forward in understanding how synthetic *a priori* knowledge is possible.<sup>6</sup> We will consider in what follows possible responses to this point, when we consider Kantian replies to Moore's objection.

Finally, then, given these premises, there seems to be considerable force in Moore's concern – where as I see it, that force primarily concerns the issue of *explanation*: Kantian idealism cannot answer its own 'transcendental' or 'how possible?' question concerning synthetic *a priori* knowledge, given that it has to assume that knowledge at a key point (when

it comes to necessary features of our minds), but without being able to explain that knowledge using the mechanism it has introduced to do so (namely transcendental idealism).

Moore would thus seem to offer a powerful internal critique of transcendental idealism, which deserves to be taken seriously.<sup>7</sup>

### **3. Response: Premise 2 and the analytic/synthetic distinction**

Having set out the problem that Moore raises, let me now consider some possible responses, where I will begin by focusing on premise 2.

Premise 2 is the claim that Kant's claim about the necessary structure of our minds is a claim to synthetic *a priori* knowledge. I think the Kantian cannot avoid treating the claim as *a priori*, as it looks unlikely to fit the profile common to the class of necessary *a posteriori* knowledge highlighted by Kripke and others, which then makes it hard to claim it is *a posteriori* at all.<sup>8</sup> But does the Kantian have to be committed to thinking that it is synthetic? Of course, the whole distinction between analytic and synthetic is far from clear cut,<sup>9</sup> but on the usual Kantian tests, how does it come out? I will consider two such tests: the logical contradiction test, and the containment test.

The first test is: if I assert *p* of *X* but I deny *q*, am I logically contradicting myself? So, if I assert of *X* that *X* is human (or 'humanly minded'), but deny that they experience the world in accordance with the forms of intuition and the understanding, is this to assert a logical contradiction, as it would be if I was asserting that a person is a married bachelor? It would seem not: for it seems an issue that we can sensibly discuss and deliberate about, as expressing disagreement on a substantive issue, rather than based on a misunderstanding of the concepts involved.

Perhaps, though, we are taking this test of contradiction too narrowly, and applying it to cases that are too simple: perhaps (as some of the logical positivists argued) there are categories of 'unobvious' analytic propositions,<sup>10</sup> which may not immediately strike us as failing the test, but which do so once considered a little more deeply, and the content of the



concepts involved are unpacked a little further. Thus, it might be claimed that when we unpack the concepts of the human mind and human experience, we will see that in fact they do logically entail the claims Kant wants to make, and we will then come to feel some logical contradiction in asserting the one without the other.

But this response in effect brings us to the second test that is relevant here, namely the ‘containment’ test: for this ‘unpacking’ metaphor suggests that in these cases, we can come to see that (as Kant puts it) the one concept is ‘contained’ in the other (cf. *CPR* A6–7/B10–11). And Kant’s test for this kind of containment is whether I need to ‘go beyond’ the one concept to get to the other: if I do not, the relation is analytic, but if I do, then it is synthetic, as I need some way of getting from the one to the other via a ‘third thing’ which thus involves more than just unpacking the terms involved.<sup>11</sup>

Now, if we adopt this test, I think it is plausible to claim that by Kant’s own lights, the connection he sees between human experience or consciousness and the forms of intuition and understanding is synthetic, because when he himself explains this connection, he also introduces a ‘third thing’ to make the link. This can be seen most clearly in the case of the Transcendental Deduction: for there, it would seem, Kant grounds the claim that ‘All human experience involves the categories’ by linking the subject and predicate of this claim via the *unity* of the ‘I think’, which shows how the one requires the other. Kant thus offers a ‘third thing’ to connect subject and predicate, in a way that suggests the connection is synthetic rather than analytic. And insofar he also brings together the forms of intuition with the forms of the understanding, he arguably in the end connects the former with human consciousness in the same way.

It would appear, then, that on the evidence of Kant’s own procedure, we are justified in taking the link between our experiences and the forms of that experience to be synthetic rather than analytic. Of course, this does not absolutely rule out a Kantian claiming that Kant

was wrong, and that he should have treated this claim as analytic, in perhaps an ‘unobvious’ way: but I find it hard to see how this could be made plausible, without compromising the whole analytic/synthetic distinction, in a way that in effect abandons the Kantian framework altogether.<sup>12</sup>

#### **4. Response: Premise 4 and the problem of explanation**

Assuming, then, that we have sufficiently defended Moore’s premise 2, what about premise 4, namely that objection that as Kant is making a synthetic *a priori* claim about the structure of our minds, he cannot offer an adequate explanation of this claim within the terms of his transcendental idealism?

I take it that the nerve of this argument is a kind of dilemma:

- a. Either Kant offers at the second-order level the same explanation as he offers at the first-order level, by appealing to a claim about the necessary structures of our minds; but then this claim itself would also need to be explained, in which case there would be a regress of such structures in a way that looks problematic in itself, and anyway would fail to provide any real explanation.
- b. Or Kant offers a different explanation of our synthetic *a priori* knowledge at this second-order level, in which case he has given us an explanation of such knowledge which does not rely on transcendental idealism, so that now the argument for this idealism looks less compelling: for in effect this was an inference to the best explanation, based on the claim that only such idealism can offer us an explanation of synthetic *a priori* knowledge – but now it is not clear why the non-idealist explanation employed at the second-order might not be used to explain such knowledge at the first-order level as well.

Let me now unpack this dilemma a little:

The trouble with the first horn (a) may be put as follows, using the terminology of a ‘third thing’ from Kant that we introduced earlier. For a synthetic *a priori* proposition, we

need some ‘third thing’ linking the subject term and the predicate term. Now, in the case of a proposition like ‘every event has a cause’, that third thing or X is said to be the schematized forms of the understanding and specifically the category of causality, which is said to be required to make experience possible. We thus now have another synthetic *a priori* proposition: ‘all human experience is structured by the category of causality’. Now, assuming (as argued above) this is also a synthetic claim, what is the ‘third thing’ here? What is it that links human experience and this category in the right way?

One answer could be: a further necessary feature of human experience, so that what makes it the case that we know that all human experience is structured by the category of causality is that we could not have experience of human beings having experience for whom this was not the case, so this second-order claim cannot be falsified within our experience any more than the first-order one can. But then of course the question arises: how do we know that? How do we know that we could not have experience of a human being who did not experience the world using the category of causality? It looks like (on this horn of the dilemma) the only answer could be the same move at the next level: but then of course we will just get a regress. And I think the trouble with the regress is that firstly, the modal claims involved look less and less plausible, or at least harder and harder to argue for or even make intelligible; and secondly, it seems to compromise the explanatory hope with which we began, which was (as Moore suggested) to ground this knowledge in something less mysterious at the second-order level than was available at the first, but we seem to just be repeating the problem.<sup>13</sup>

In response, then, we might try the second horn (b), but this seems equally problematic. For example, we might claim that we know all human beings must experience the world in causal terms, because without this their experience would become disunified to such a degree that they could not be experiencing selves at all. So here, the ‘third thing’

might be said to be ‘the unity of the self’ which links being a human experiencer or mind with employing the categories in one’s experience. But now, on Kant’s account, what is the explanation for our knowledge of this synthetic *a priori* claim? How does Kant know this to be the case? One answer might be along the lines suggested above, but as we have seen, this leads into difficulties. To avoid this, another approach might be: he knows this, because he can just see the various metaphysical relations between these concepts, as a kind of transcendental argument from the one to the other – without the unity imposed on our experience through the use of various synthesising categories, experience would be impossible as it could not fall within the bounds of a unified self having that experience.

But then the question arises: how does he know about these metaphysical relations? If the answer is: he just has a kind of metaphysical knowledge here, in tracing out various necessary connections which explain how the one requires the other, then he has helped himself to a kind of metaphysical knowledge that does not have any explanation in the terms used by transcendental idealism to explain such knowledge. And if that answer works here, why is transcendental idealism needed at all to explain how such knowledge is possible? Moore himself puts this point as follows:

And if, in spite of the fact that it does not follow [from experience], you are able to know that all men’s minds always have and always will act in this way; why should you not be able to know that *all* triangles, even if triangles are not merely appearances produced by the action of your mind, must have their angles equal to two right angles? (Moore 1953: 154)

But if, on the other hand, Kant deploys the explanatory machinery of transcendental idealism, and so grounds this metaphysical knowledge in the necessary conditions for our experience or mindedness, he would appear to be back onto the first horn of the dilemma, and the problems that seemed to entail.

At this point, one response to the problems posed by this dilemma might be to try to make a distinction between the two levels, and claim there is a significant epistemic difference between them. Namely, our knowledge of the necessary and universal features of the world is one thing, and our knowledge of the necessary and universal structures of our mind is another. For, it could be argued that while the former needs transcendental idealism to explain it, the latter does not, as in the case of our own minds, we have access to those features in some other way, perhaps through some direct metaphysical insight. This point could be put in the terms of Kant's famous letter to Herz of 1772, in which some of the key motivations and thinking behind the 'transcendental turn' were initially sketched out: while an account is needed of how our faculty of understanding achieves conformity with things which avoids just assuming some *deus ex machina* to co-ordinate them (Kant *Corr*, 10: 131), no such account is needed of how it is that we can grasp the necessary structures of our minds, as precisely such structures are not 'external' to us in the same way, so that we can know them directly.

However, while this is a possible response, I think it can be challenged. For, what reason have we got, other than a lingering Cartesianism, to think that our own minds are more metaphysically transparent to us than anything else? As Moore argues, if these claims are based on our experience of how minds work, then they are no more able to give us necessity than any inductive argument:

From no amount of experience, which you may have had as to how your mind acted in the past, will it absolutely follow that it ever will do so again, or that it ever has done so, except in the instances which you have actually observed; nor will it follow that any other man's mind ever has acted or ever will act in the way in which yours has. (Moore 1953: 154)

On the other hand, if these claims are not based on experience, the question arises here precisely as it does with respect to the world outside us: how could we possibly come to know these claims are true? This question concerning how synthetic *a priori* knowledge of our minds is possible seems no less pressing than how synthetic *a priori* knowledge of the world outside those minds is possible.<sup>14</sup> And if Kant has an answer to this question, it seems it would have to avoid an appeal to transcendental idealism if it is to escape the regress outlined above. But then, if transcendental idealism is not required to explain synthetic *a priori* knowledge when it comes to the structure of our minds, why should it be required to explain such knowledge regarding the world outside our minds – in which case, how has Kant ruled out metaphysical knowledge of a transcendental realist kind?

Another way to put this point has been suggested by Dustin McWherter, who frames it in terms of transcendental arguments, and the suggestion that Kant in effect offers a transcendental argument that forms of intuition and the categories are necessary conditions for the possibility of human experience. But then the question arises: how does Kant know this argument holds? Either his knowledge comes through knowing further necessary conditions for the possibility of experience; but then a regress threatens. Or his knowledge does not depend on knowing about such conditions but instead on knowing certain metaphysical facts about our minds; but then we have a realist transcendental argument, so that it appears that we do not have to confine our modal claims within the conditions of experience, and so can extend them to things in themselves.<sup>15</sup>

I will now consider two responses to this Moorean challenge. The first of these suggests that Kant can meet the challenge by appeal to his account of transcendental reflection (section 4.1); and the second response suggests he can do so by appeal to inference to the best explanation rather than a transcendental argument (4.2).

#### ***4.1 Transcendental psychology and transcendental reflection***

This first response in effect claims that Kant himself provides grounds for treating our knowledge of our minds and our knowledge of the world outside our minds as different; this difference can then be used to explain how synthetic *a priori* knowledge of our minds might come about without a Copernican explanation, while also showing why our synthetic *a priori* knowledge of the world is more challenging, and therefore does require such an explanation. The justification for attributing this approach to Kant could be said to come from the way in which Kant himself seems to put self-knowledge of the sort found in transcendental psychology on a different footing from our knowledge of the world – and this is perhaps why it never occurred to him to think the problem of synthetic *a priori* knowledge that arises for the latter arises in the same way for the former, and thus why he may not have felt he really needed to grapple with the Moorean problem in the first place.

Now, it could certainly be said that a central point of Kant's Paralogisms and his criticism there of rational psychology is to argue that the transcendental subject is not to be confused with an object in the world, and thus our investigation of it must be fundamentally different. The question still arises, however, how we are to make claims about this subject, as a matter of transcendental rather than rational psychology. Here is not the place to plumb the full depths of this transcendental psychology and all that it implies, but one aspect of its method that has attracted quite a lot of (mostly positive) attention recently, and which might seem particularly relevant to our issue here is Kant's notion of 'transcendental reflection', and what that seems to involve. For, it could be argued, Kant precisely introduces this notion in his Appendix to the Transcendental Analytic ('On the Amphiboly of the Concepts of Reflection Through the Confusion of the Empirical Use of the Understanding with the Transcendental') in order to give some epistemic account of how he has been proceeding in the Analytic, and thus making the various synthetic *a priori* claims about the human mind that he has been presenting in this key section of the *Critique*. Thus, for example, Kenneth

Westphal has argued that by taking what Kant says about transcendental reflection more seriously, we can escape the kind of dilemma posed for transcendental arguments that we discussed above (see Westphal 2004: 16–17).<sup>16</sup>

However, while it is clear that Kant does give transcendental reflection a role that might be used to make sense of aspects of his transcendental psychology, it is not initially obvious that he uses it to explain the kinds of synthetic *a priori* claims about the mind that interest us. Kant's own definition of transcendental reflection is as follows: 'The action through which I make the comparison of representations in general with the cognitive power in which they are situated, and through which I distinguish whether they are to be compared to one another as belonging to the pure understanding or to pure intuition, I call *transcendental reflection*' (CPR A261/B317). Kant thus states that the reflection is primarily aimed at answering the fundamental question: 'In which cognitive faculties do [our representations] belong together? Is it the understanding or is it the senses before which they are connected or compared?' (A260/B316). Now undoubtedly, this is an important issue for transcendental psychology in general, but it is not clear it is designed to address the issue that Moore raises, namely to give us the kind of modal knowledge of the mind that seems required to make sense of the Copernican revolution. It is true that Kant says that 'This transcendental reflection is a duty from which no one can escape if he would judge about things *a priori*' (A263/B319). But it is not obvious that Kant is claiming here that transcendental reflection is how we *do* judge about things *a priori*, rather than just claiming that unless we correctly identify the source and status of representations and so handle their 'transcendental placement' accurately (cf. A268/B325), we will be misled concerning the *a priori* claims we make about them, as Leibniz for example is said to be (cf. A267/B323, A271–5/B327–32) – but this is in its way just another statement of the Copernican revolution, rather than an explanation of it, and so cannot help in addressing the Moorean challenge.



Nonetheless, this section of the *Critique* is often taken to be suggestive in a more general manner, for the emphasis it places on the form/matter distinction,<sup>17</sup> and also for the parallels it draws between logical and transcendental reflection, where the *a priori* status of both logic and transcendental psychology is said to be based on their focus on form, which reflection is then said to be able to reveal. In an influential article, Smit makes this parallel clear when he writes:

Thus, just as logical reflection must distinguish the logical form of a thought from its matter in order to determine the logical relation which concepts bear to each other in a judgment, so too, transcendental reflection must distinguish the form and matter of our cognition of an object in order to determine the real relations which the representations making up the manifold of an intuition bear to each other in a cognition of an object. (Smit 1999: 216)

On this basis, Smit then goes on to apparently address the issue that concerns us here:

We thereby isolate the form of our cognition of objects, and with it, certain (descriptively) necessary laws which obtain among all possible objects of our cognition (e.g., the principles of pure understanding) because they are constitutive of the form of our cognition of others. And, we are capable of isolating this form priori, so as to cognize universal and necessary laws governing all possible objects of our cognition, because all our cognition of objects, including that which constitutes our experience, has transcendental reflection as an essential component. (ibid)

In this way, then, Kant's conception of transcendental reflection and its ability to give us access to the forms of cognition may seem to explain why synthetic *a priori* knowledge is possible in this arena but not correspondingly possible for the kind of knowledge claimed by the metaphysician, as the latter cannot appeal to the capacity for transcendental reflection to

ground this knowledge, as this kind of *self-conscious* reflection only extends to the mind and not the world.

However, it appears to me that Moore might legitimately remain dissatisfied with this response: for this still seems to rest on what again looks like a synthetic *a priori* claim, namely that to abstract from the matter of cognition and so focus on form thereby gives us access to the ‘universal and necessary laws governing all cognition’, as it does not seem analytic that the form we have hereby identified has these kinds of modal properties. How does any process of abstraction in itself warrant the claim, for example, that such forms are ‘constitutive’ of cognition, and that they are also and thereby necessary and universal forms of cognition? We thereby once again appear to come back to a synthetic *a priori* judgement concerning the modal properties of this form which an appeal to transcendental reflection itself seems unable to explain. Taken on its own, therefore, the appeal to transcendental reflection would not seem to bridge the explanatory gap concerning our ability to acquire modal knowledge concerning the transcendental subject, and without this knowledge the Copernican revolution still looks under threat.

#### ***4.2 Transcendental psychology and inference to the best explanation***

I now turn to consider a second response to the Moorean challenge concerning this fourth premise, and how we can acquire synthetic *a priori* knowledge of our minds. On this response, it can be argued that Kant is not using a transcendental argument to *establish* the necessary structures of our minds, but is instead offering an *inference to the best explanation* (IBE). That is, *from* the fact that we have synthetic *a priori* knowledge, he is inferring to the best explanation of that fact, namely that our minds have certain necessary structures to which experience must conform.<sup>18</sup> So Kant could then respond to Moore that this IBE enables him to make the claim that the mind has certain necessary structures, where, as an abductive argument, we can explain how it gives us knowledge without generating a regress as such arguments do not rest on prior synthetic *a priori* knowledge; and he could respond to

McWherter along similar lines, arguing that the regress can be avoided, as he is not offering a transcendental argument for the claim that our minds must have a particular structure, but an IBE instead, as the best explanation for how synthetic *a priori* knowledge is possible.

An obvious reply to this response might be that it would seem to generate its own sub-dilemma: on the one hand, if using IBE on its own is sufficient grounds on which to make synthetic *a priori* claims about the necessary structure of the human mind, the realist might wonder why we cannot hope to use IBE to make synthetic *a priori* claims elsewhere, about the world beyond such minds; but on the other hand, if IBE is not a sufficient ground to make such claims about the necessary structure of the human mind, we would then not have escaped the main dilemma. However, here I think Kant could respond: his IBE from our synthetic *a priori* knowledge takes us to transcendental idealism, as the best way to explain that knowledge – but then, transcendental idealism shows our knowledge to be limited in ways that show we cannot expect IBE to work for a realist metaphysics, as we then do not know what a ‘best explanation’ outside the boundaries of our experience could mean.

Nonetheless, while this reply might seem adequate as far as it goes, I think Kant’s critic can still raise other significant objections to the IBE response. First, it relies on taking synthetic *a priori* knowledge for granted, as the given that needs to be explained. One obvious difficulty is that while Kant might be confident that we have such knowledge in arithmetic and geometry, as he took the validity of Euclidean geometry to be unassailable for example, we cannot be so sanguine. But even more problematically, while Kant might think he is entitled to take synthetic *a priori* knowledge in arithmetic and geometry for granted,<sup>19</sup> arguably even he isn’t that confident when it comes to knowledge that looks more metaphysical, such as that ‘every event has a cause’. Indeed, it appears that the role of the Transcendental Deduction is precisely to *ground* or *vindicate* such knowledge, not simply to take it as a given<sup>20</sup> – and it is in the process of doing this that the dilemma then seems to

arise, as he uses his appeal to the necessary conditions of experience to warrant such knowledge, rather than merely rely on it to make a further argument.<sup>21</sup> Second, if Kant is operating with an IBE and thus an abductive argument, this is of course a weaker form of argument than a transcendental argument, and can deliver only degrees of plausibility but not certainty; but then it is not clear that it would not leave some room for doubt concerning the status of transcendental idealism, thus leaving the door open for realist synthetic *a priori* claims still to be made. And finally, in response to this worry, the Kantian might claim that his IBE is stronger than the usual form of such argument, as really he is claiming that transcendental idealism is the *only* possible explanation of our synthetic *a priori* knowledge, as all other explanations would violate the conditions under which our minds must operate – but then to strengthen the IBE in this manner based on a modal claim of this sort would in effect bring us back to the original dilemma, and turn it into a form of transcendental argument again.<sup>22</sup>

### 5. Response: Premise 3 and the limits of explanation

I now want to consider another premise of the Moorean argument above, namely the third premise, which claims that Kant does not explain how this knowledge of the necessary structure of our minds is possible. As I noted above, while as a textual claim this seems true, I now want to briefly explore whether the Kantian might adopt a strategy that Kant is happy to use elsewhere, and so perhaps could be used by the Kantian here, namely to claim that there are limits to human understanding and our capacities for explanation, and that at some junctures of a philosophical inquiry, this just has to be accepted. For example, Kant famously argues in the *Groundwork* that ‘reason would overstep all its limits if it ventured to *explain how pure reason can be practical*, which would be the same as explaining *how freedom is possible*’ (*G*, 4: 458–9), so that while we can know (on practical grounds) that we *are* free, we cannot really understand how. Likewise, it could perhaps be said, we know that the

human mind *does* have universal and necessary structures, but this is a kind of knowledge we *cannot explain* for reasons we have discussed – it just has to be accepted as being the case.

Now, some may find this kind of Kantian move objectionable in itself – as a kind of ‘cop out’ which means you can just duck hard philosophical questions by feigning incapacity.

However, I do not object to this kind of move in principle – but to be credible, it has to be properly warranted, and cannot be used just in order to get yourself off the hook. In the places where Kant generally uses it, he does offer a principled argument for our incapacity, usually based on an appeal to the structures of transcendental idealism – as he does in the case of the *Groundwork* above, in playing on the kind of dualism between freedom and necessity this implies, so that (as Kant nicely puts it) ‘although we do not comprehend the unconditional practical necessity of the moral imperative, we do comprehend its *incomprehensibility*’ (*G*, 4: 463). However, the question in the case we are considering is whether the Kantian is just claiming brute incomprehensibility, or has some way of helping us ‘comprehend’ it – where the latter seems required, if this is not just to be an ad hoc way of simply ducking the problem in an unsatisfying manner.

It seems to me that the difficulty in this case is that there are not the usual Kantian grounds for drawing a limit to our capacities here: Given that we can know that the human mind has such necessary structures, what grounds could the Kantian offer for making plausible our incapacity to understand this necessity? Perhaps the answer could be: because here we are coming to know something of how the noumenal world works, and we can never explain how we have such knowledge, even though here we must assume we do. But if the Kantian is allowed to play this card at this level, why cannot his opponent play the same card at the first level, and say that our metaphysical knowledge also relates to the noumenal realm, and therefore it too defies explanation for the same reason, but that this gives us no reason to question it, any more than it does for Kant himself? Moreover, in other cases that Kant

discusses, the knowledge we find hard to explain comes from practical reason, where it is then perhaps plausible to think (given Kant's position) that this outstrips the capacities of theoretical reason to fully comprehend it. But in this case, the knowledge is itself based on theoretical reason, which also makes this move less plausible: to give us theoretical knowledge, does not theoretical reason need to be able to understand how that knowledge is possible?

However, a different strategy relating to explanatory issues is suggested by Colin Marshall (Marshall 2014). He does not argue that Kant might hold that some explanatory matters are beyond our understanding, but rather that Kant might hold that some synthetic *a priori* claims are 'explanatorily basic', and so require no explanation in the first place, where such claims about the structure of the mind are said to fall into this category for Kant. Nonetheless, as Marshall recognizes, this assertion of basicness must be earned, where he suggests that Kant held this view because he took such knowledge of our mind to be intrinsically less mysterious than synthetic *a priori* knowledge about the world, as the former is accessible to what Marshall calls "short-range rationalist reflection" in a way that the latter is not (p. 564). However, in claiming Kant adopts this approach, Marshall largely appeals to 'transcendental reflection' along the lines discussed in section 4.1 above; but as argued there, it is not clear that the Moorean would be satisfied with this response. Moreover, Marshall admits that there are some potential costs to attributing this kind of rationalism to Kant, as 'many philosophers would deny we have even such a modest capacity (e.g. Hume), or if they accepted it, might conclude that it yields different conclusions from what Kant thought (e.g. Hegel)' (p. 572); however, Marshall thinks these costs must be paid, as he does not think there are any other interpretative options, once we reject other approaches, such as holding that such claims are analytic, not synthetic, or that Kant can employ an inference to the best explanation. As discussed above, I agree with Marshall over his negative arguments

against these approaches; but now I want to introduce a further option that Marshall does not consider and suggest it is more promising than the proposals he rejects, and also thus more promising than his own.

### **6. Response: Premise 1 and the metaphysics vs the normativity of mind**

Having looked at all the other premises of the Moorean argument, we only have one final premise to examine, namely premise 1: ‘Kant explains how synthetic *a priori* knowledge is possible by appealing to the necessary structures of our minds’. This seemed unexceptionable, as it seemed like a natural way to characterize Kant’s Copernican revolution: to explain the necessary structure of our experience, we turn to see how this experience is grounded in the necessary structure of our minds, thus moving (as it were) from the metaphysics of the world to the metaphysics of our minds. Thus, Moore himself talks about Kant making a claim about ‘the constitution of our minds’, while other critics speak of knowing *a priori* ‘how the mind operates’ or of facts concerning our ‘cognitive constitution’.<sup>23</sup> But then, of course, the Moorean challenge is to question whether making metaphysical claims about how our minds must work is any easier to understand than making metaphysical claims about the world outside our minds

My final suggestion on Kant’s behalf, therefore, is to question whether this innocuous looking first premise is the right way to characterize the Copernican turn after all, and in particular whether that should be understood as offering a kind of ‘metaphysics of mind’ – for of course, once this is accepted, it then gets hard to see why the realist cannot go back to offering their ‘metaphysics of the world’, for the reason we have explored. The point of this final response is therefore to question whether Kant is indeed offering a ‘metaphysics of the mind’, namely making synthetic *a priori* claims concerning the necessary functioning of our minds; for then if this is not what he is doing, the equivalence claim will fail.

But, if Kant is not conducting a metaphysics of the mind, what *is* he doing – indeed, what else *could* he be doing? To see how there might be other options here, it is important to

focus not on the metaphysical questions we might raise regarding the mind, but rather *normative* questions: for, if it is appropriate to think of the mind as involving knowledge, justification, understanding, judgement, reasoning and so on, it is appropriate to think of it in normative terms – that is, as having to meet various norms in order for it to possess knowledge, justification, understanding and so on. Thus, to illustrate the distinction I have in view, we might ask these two questions regarding memory, for example: first, ‘would a person whose memory became radically unreliable lose their experience of time passing, as they would then lack the information required to make this experience possible?’, which I take to be a question in the metaphysics of mind, which we might try to investigate in an *a priori* manner; and second, ‘would a person whose memory became radically unreliable be able to acquire knowledge of past events?’, which I take to be a question in the normativity of mind, as it asks how epistemic norms could be met by such a being.

Now, it is admittedly the case that in the Transcendental Aesthetic it can seem that Kant is engaged in a metaphysics of mind and thus a transcendental psychology, as a way of addressing the problem of synthetic *a priori* knowledge. Thus, for example, it is this section of the *Critique* that leads Bertrand Russell to adopt his famous (or notorious) ‘blue spectacles’ analogy for space and time as the forms of intuition, which treats them as akin to ‘spectacles in the mind’ – so just as ‘if you always wore blue spectacles, you could be sure of seeing everything blue’, so likewise ‘since you always wear spatial spectacles in your mind, you are always sure of seeing everything in space’.<sup>24</sup> It is precisely this picture that invites the Moorean challenge that we have been discussing. However, it becomes easier to see how a more normative approach might work when Kant develops his more complete picture and moves from intuition to judgement, and thus into the Transcendental Analytic, where we might for example focus on the Second Analogy and Kant’s claims there about causality.



It is not possible here to dwell on this in any detail, but the core idea could be expressed as follows: If we did not apply the concept of cause to a sequence of representations and thus subsume them under a rule, we would have no justification for distinguishing them from subjective experiences and thinking of them as having an objective temporal order instead – and so could not take them to be events at all. Kant makes this normative point clear when he writes:

If, therefore, we experience that something happens, then we always presuppose that something else precedes it, which it follows in accordance with a rule. For without this *I would not say of the object that it follows (daß es folge)*, since the mere sequence (*bloße Folge*) of my apprehension, if it is not, by means of a rule, determined in relation to something preceding, *does not justify any sequence in the object (keine Folge im Objecte berechtigt)*. (CPR A195/B240, my emphases)

In this case, therefore, the claim that every event has a cause is derived from consideration of what *entitles* or *justifies* us in taking our apprehension to be of an event in the first place, which is an entitlement we cannot possess unless we are prepared to apply the concept of cause to our experience: unless we apply the concept of cause, we would lack the criteria by which to treat it as an experience of an event rather than as part of a subjective sequence, and so could not claim to have any knowledge of it *as* an event to begin with. As Kant immediately goes on to note after the passage above, this amounts to an instance of the Copernican turn: rather than claiming that we first experience events, and then infer by induction that all events have causes, which makes the latter claim contingent, the normative basis of the argument above gives us grounds for taking the causal principle that every event has a cause to be necessary and knowable *a priori*, based not on how the world is in itself, but on normative constraints on how our cognition works.<sup>25</sup>

Likewise, in the First Analogy, Kant arguably also takes a normative approach, in arguing that to know that a substance has come to be or ceased to be, this transition has to involve a change in what is rather than being absolute; for if it were absolute (such as a creation *ex nihilo*) we would not be able to judge it to be an event in time, as this precisely requires relating it to what existed previously, for this is the only way we can put our subjective temporal experiences into an objective temporal order. For, to judge on the basis of your experiencing A now that A has just come into existence *ex nihilo*, you would have to have some grounds for placing A's coming into being at that temporal point, rather than A having existed previously and you having only just noticed it, and so you must be entitled to make the judgement that at some point of time it did not exist in any form; but in saying that it is unrelated to anything that existed before it, you could not be warranted to make any judgement about its place in time, as this precisely requires relating it to what existed previously, for this is the only way we have for legitimately putting our subjective temporal experiences into an objective temporal order. Thus, we can never be in the position to judge with any justification that a substance has come into existence *ex nihilo*; rather, if this appears to be the case, we should question our experience, not endorse it.<sup>26</sup>

Now I am not in any way trying to claim that the details of these arguments are unproblematic. But my point here is a more general, strategic one: because Kant's argument revolves around a normative issue, it does not rest simply on a synthetic *a priori* claim concerning the metaphysics of the mind, and so breaks the equivalence that Kant's critics exploit in objecting to his account. For, on this reading, Kant is basing his explanation of synthetic *a priori* knowledge on normative claims about what justifies us in making judgements using certain concepts such as 'event' or 'coming to be', from which then synthetic *a priori* claims about our experience follow.<sup>27</sup> It turns out, then, that the 'third thing' which makes such first-order synthetic *a priori* propositions possible must be

understood in normative rather than modal terms. But then the *a priori* metaphysician is not, and indeed cannot be, making an equivalent move, for the metaphysical claims about the world they want to make are not derived from any normative pre-conditions, parallel to the way in which Kant exploits the conditions necessary for knowledge, understanding, judgement and the like in order to arrive at his modal claims regarding the nature of our experience. Instead, metaphysicians are taking themselves to have some insight into the necessary structure of what there is, which looked hard to challenge if Kantian were themselves claiming insight into the necessary structure of mind as a component of what there is; but if we take the normative approach outlined above, this parallel is broken as Kant's claims concern the mind as a normative entity rather than a metaphysical one. Thus as a result, the Kantian and the metaphysical methods can no longer be treated as equivalent, making it harder for the metaphysician to claim that however the Kantian explains synthetic *a priori* knowledge with regard to the mind, they can do the same with respect to the rest of what there is – for Kant's approach turns out to depend on the normative features of the former that do not apply to the latter.

Moreover, it is arguably an advantage of this approach that when put in a normative rather than a modal context, some of the options we considered above now look more promising, particularly the issue of analyticity and of transcendental reflection.

On the first issue, Kant's discussion of the categorical imperative in the *Groundwork* shows that Kant did not think that moral norms governing practical reason are analytic (see *G*, 4: 417–20, 445), and as a result he allowed that to explain it requires 'special and arduous toil' (4: 420). Nonetheless, he also seems to treat this as a rather special case, and his way of arguing that the instrumental norm is analytic could be extended to also argue that the normative claims about the mind that concern us here are likewise analytic, which was harder when that analyticity was meant to cover modal claims about the mind, not normative ones

(see section 3). That is, when the ‘must’ in question becomes a normative ‘must’ rather than a modal one, it becomes easier to see how it could be claimed to be required for rational judgement which is itself a normative matter, just as the instrumental norm is required for rational agency.

And relatedly, on the second issue, insofar as in various writings on logic Kant treats logical reflection as reflection on the normative laws of thinking in general (cf. Smit 1999: 214–5),<sup>28</sup> once the issues are reconceived in normative terms, then likewise the form of cognition identified by transcendental reflection can be understood to be reflection on the normative rules of thinking about representations, and so be more effectively used to address these questions than seemed possible above (section 4.1), before these normative considerations were made central. Thus, if the proposals I have made in this article are deemed worthwhile, a consideration of these options could be a fruitful next stage for the discussion to take.

But my more limited claim here is just that however we answer these questions, Kant’s arguments about the mind that are required to show how synthetic *a priori* knowledge of the world is possible do not involve a kind of modal knowledge but a normative knowledge; and while as we have seen, the metaphysician may reasonably wonder why if we can have modal knowledge of the mind, we cannot equally have modal knowledge of the world in the same way, there is no equivalent move from *normative* knowledge of the mind to modal knowledge of the world, as the two are no longer on a par. Likewise, while it was difficult on other strategies to claim that self-knowledge and worldly knowledge are significantly different when both were taken to involve modal claims, this becomes easier when the former concerns what is normative and the latter concerns the modal. Of course, it is then perhaps a challenge to see how normative claims about the mind can yield modal claims about the world: but then it is this challenge that transcendental idealism is designed to

address as the examples from the two Analogies above briefly illustrate, but now in a way that is immune to the Moorean objection. To this extent, therefore, what appeared to be the Achilles' heel of transcendental idealism turns out to be adequately protected after all.<sup>29</sup>

## Notes

<sup>1</sup> As McWherter (2016: 201–2) notes, there have been brief replies in Bird 2006 and Allais 2010, while a fuller response by Marshall 2014 is discussed further below. A wider but related issue concerning the stability of Kant's transcendental psychology has been much more frequently considered, namely whether Kant's claims about the latter overstep the bounds of inquiry set by his transcendental idealism. As we shall see below, some of the responses considered in this debate, particularly the appeal to 'transcendental reflection', might also be said to apply to the issue here, although the two issues are not the same: one might think transcendental psychology can be defended in some respects, without it following that it is unproblematic for this psychology to establish what Kant requires for his explanation of synthetic *a priori* knowledge to go through.

<sup>2</sup> Cf. CPR B19: 'The real problem of pure reason is now contained in the question: *How are synthetic judgments a priori possible?*' References to Kant are given in the standard form, using A and B to refer to the first and second editions of the *Critique of Pure Reason* respectively, and to the pagination of the Akademie edition for other works. Abbreviations are used as follows: *Corr* = *Correspondence* (1999); *CPR* = *Critique of Pure Reason* (1787/1998); *G* = *Groundwork for the Metaphysics of Morals* (1785/2019); *P* = *Prolegomena to Any Future Metaphysics* (1783/1997).

<sup>3</sup> Cf. *P*, 4: 276.

<sup>4</sup> For some further discussion, see Stern 2009: 23–4.

<sup>5</sup> This text is based on lectures Moore gave at Morely College, London in the winter of 1910–11, and thus not long after he also wrote ‘Kant’s Idealism’ (Moore 1903–4) in which he first expressed these criticisms.

<sup>6</sup> Cf. Nietzsche 1886/1966, §11, 18–19 (cited Marshall 2014: 562, n. 43): “How are synthetic judgments *a priori* possible?” Kant asks himself – and what really is his answer? “*By virtue of a faculty (Vermöge eines Vermögens)*” – but unfortunately not in five words ... “By virtue of a faculty” – [Kant] had said, or at least meant. But is that – an answer? An explanation? Or is it not rather merely a repetition of the question? How does opium induce sleep? “By virtue of a faculty”, namely the *virtus dormativa*, replies the doctor in Molière.’

<sup>7</sup> As I have said, other subsequent critics have picked up on these concerns, some associating them with Moore and some not: see for example Bennett 1966: 16–18, Van Cleve 1999: 37–43, Forster 2008: 61–70, Millican 2017: 59. Other reference to the literature on this issue are given below.

<sup>8</sup> This more *a posteriori* option has nonetheless been pursued by Derk Pereboom (e.g. 1990) and Patricia Kitcher (e.g. 1994), who argue that Kant’s transcendental assertions about the mind might be given some sort of empirical justification; but as Marshall 2014: 558–9 points out, it is hard to see how any such justification can be sufficient to support the kind of modal claims that seem to be required, for reasons Kant himself famously asserts at *CPR* A1/B3: ‘[Experience] tells us, to be sure, what is, but never that it must necessarily be thus and not otherwise’. For Moore’s own worries about this approach, see Moore 1953: 154, which is cited below. I will therefore set this option aside without further comment.

<sup>9</sup> It might also be said not to be exhaustive on some definitions of ‘synthetic’, but as Marshall 2014: 553 points out, in addition to some textual evidence suggesting Kant thought it was, the problem with *a priori* knowledge arguably arises as soon as we move beyond just the analytic cases, so to that extent this issue is moot.

<sup>10</sup> Cf. also Sellars's discussion of 'illuminating' analytic truths, which O'Shea has used to defend an analytic approach to Kant here: see O'Shea 2019.

<sup>11</sup> Cf. *CPR* A155/B194: 'If it is thus conceded that one must go beyond a given concept in order to compare it synthetically with another, then a third thing is necessary in which alone the synthesis of two concepts can originate.'

<sup>12</sup> It could be argued that there is in fact textual evidence for the analytic approach at *CPR* A65–6/B90–1, where Kant introduces his 'Transcendental Analytic'; but I would agree with Marshall that here 'Kant contrasts his project of analyzing the faculty of understanding with the more common project of analyzing concepts' (Marshall 2014: 577, n. 29). For another example of where Kant seems to use analysis in this different sense, see A842/B870, when he compares his procedure to the analytical method of the chemist.

<sup>13</sup> Cf. BonJour 1998: 24: '[H]ow would the synthetic *a priori* status of this new proposition [*P*\*: within the bounds of experience, *P* is the case] be accounted for? To offer the same account that was originally offered for *P* would in effect require Kant to say that the mind so operates as to make it true that: the mind so operates as to make it true that *P*\*. It is doubtful that this new claim even makes sense, but even if it does, the same question can be raised about it, and so on *ad infinitum*, thus generating a regress that is clearly vicious.'

<sup>14</sup> Forster 2008: 133, n. 13 suggests that in fact for Kant it is *more* pressing, given the epistemic limits Kant puts on philosophical inquiry, and that the transcendental self is not part of the phenomenal world: 'Kant's suggestion here [*CPR*, Axiv] of a Cartesian self-transparency of the mind is indefensible within the framework of the critical philosophy, at least if understood as a doctrine about the mind and its activities *in themselves*, as it would have to be in order to explain our knowledge of the thesis of transcendental idealism.' But for an earlier defence of Kant on this score, see Smit 1999.

<sup>15</sup> Cf. McWherter 2016: 203: ‘In other words, it is not enough to simply appeal to Kant’s transcendental arguments as the means by which transcendental subjectivity is cognized *a priori*. Instead, you also have to show that things in themselves cannot likewise be cognized *a priori* as a result of transcendental arguments. For if they can, Kant loses his reason for rejecting transcendental realism, since things in themselves would be cognizable *a priori* after all.’ McWherter uses the argument from Moore to suggest that it is hard for the Kantian to resist the claim, made by Ralph Walker, that transcendental realists can also deploy transcendental arguments.

<sup>16</sup> Westphal comments that ‘Kant’s account of transcendental reflection, like his name for it, are conspicuously rare, almost absent, from Kant scholarship’ (Westphal 2004: 2). He overlooks Smit 1999, while more recently the issue has had quite wide discussion, e.g. Marshall 2014, Merritt 2015, Balanovskiy 2018, De Boer 2020: chapter 7, while it is also important to Schafer’s recent ‘capacity-first’ approach (Schafer 2021).

<sup>17</sup> ‘[Matter and form] are two concepts that ground all other reflection, so inseparably are they bound up with every use of the understanding’ (*CPR* A266/B322).

<sup>18</sup> See e.g. *CPR* A48–9/B65–6, B166–8 and B41: ‘Thus our explanation alone makes the *possibility* of geometry as a synthetic *a priori* cognition comprehensible.’

<sup>19</sup> Though even this is methodologically problematic with respect to the *Critique of Pure Reason*, as it would muddy the distinction Kant himself draws between the synthetic method of the *Critique* which Kant claims ‘takes no foundation as given except reason itself’, and that of the *Prolegomena* which he says proceeds analytically by relying ‘on something already known to be dependable’, such as the synthetic *a priori* knowledge of pure mathematics and pure natural science (*P*, 4: 274–5).

<sup>20</sup> Cf. *CPR* A87–92/B119–24.



<sup>21</sup> Cf. *CPR* Bxviii: ‘This [Copernican] experiment succeeds as well as we could wish ... For after this alteration in our way of thinking we can very well explain the possibility of cognition *a priori*, and what is still more, we can provide satisfactory proofs of the laws that are the *a priori* ground of nature, as the sum total of objects of experience – which were both impossible according to the earlier way of proceeding.’

<sup>22</sup> In considering this debate, I am leaving aside another possible Kantian response: that Kant can offer a different IBE for transcendental idealism based on the dialectic of reason, with the latter providing a ‘checkup’ (*Gegenprobe*) or further confirmation of the transcendental turn (*CPR* Bxx). It would take us too far afield to consider the adequacy of this argument, and anyway our interest was in seeing whether the problem of synthetic *a priori* knowledge could motivate transcendental idealism on its own.

<sup>23</sup> BonJour 1998: 24, Van Cleve 1999: 38.

<sup>24</sup> Russell 1961: 680.

<sup>25</sup> The normative aspect of this approach is emphasised by Paul Guyer in his reading of the Second Analogy: ‘Note Kant’s remark [at Kant 1787/1998, A193/B238] that the function of such a rule is that of *justifying me in saying* (or judging) of an empirical object that it contains an objective succession. Here – and almost nowhere else – Kant precisely delineates just what it *means* to call a principle such as that of causation a principle of the possibility of experience. It is not to say that such a principle is one that constitutes an empirical object in any ontological sense, nor that it is one which is somehow a psychological precondition for the occurrence of a representation, even a propositional representation of – a belief about – an object. Rather, to call a principle a condition of the possibility of experience is to say no more and no less than that it is a necessary condition for the *justification, verification, or confirmation* of the judgments about empirical objects that we make on the basis of our representation of them – to whatever degree of confirmation they actually admit. It is in this

sense, that of constituting the framework for our epistemic practices of judging about objects, rather than that of constituting either the objects themselves or the psychological processes by which we come to have images of or beliefs about them, that Kant means his statement that the business of the understanding “is not that it makes our representations of objects distinct, but that it makes the representation of an object possible at all” (A199/B244)’ (Guyer 1987: 245–6). For a more recent reading which also focuses on a normative account of the Second Analogy, see Hutton 2019, who puts this approach in a wider context which includes Onora O’Neill, Henry Allison, John McDowell, Konstantin Pollok and others (pp. 593–4). As far as I know, none of these proponents of a normative approach have used it to address the issue I am dealing with in this article.

<sup>26</sup> Cf. *CPR* A188/B231. As Jay Rosenberg notes, there are some echoes here of Descartes’s position in the Sixth Meditation, which is also put in normative terms, concerning what we have reason to believe: ‘[If] someone, while I was awake, quite suddenly appeared to me and disappeared as fast as do the images which I see in sleep, so that I could not know from whence the form came nor whither it went, it would not be without reason that I should deem it a specter or a phantom formed by my brain [and similar to those which I form in sleep], rather than a real man’ (Descartes 1964–76: VII, 89–90; cited Rosenberg 2005: 211–12).

<sup>27</sup> I thus agree with Hutton when he argues against Pollok that Kant is not treating propositions like ‘every event has a cause’ as synthetic *a priori* because they are *themselves* norms for cognition, but that he explains the synthetic *a priori* status of these propositions based on normative claims of the sort outlined above: see Hutton 2019: 606, where he is criticising Pollok’s assertion (2017: 2) that ‘synthetic judgments a priori must be acknowledged as the fundamental norms for our mathematical and empirical cognitions’.

<sup>28</sup> There has of course been some controversy on this issue, provoked by Tolley 2006, but that cannot be discussed further here. For one response, see Hutton 2019: §6.

<sup>29</sup> Earlier versions of this paper were presented to conferences at the University of Cologne and at Royal Holloway, University of London; I am grateful to audiences for their comments, and also for comments separately by Fraser MacBride, Joe Saunders and Marcus Willaschek. I am particularly grateful for extensive comments from Adrian Moore and Jessica Leech, and also to Paul Franks for first suggesting to me that a normative approach to these issues might be fruitful. Finally, the two anonymous referees for this journal also provided very helpful sets of comments which greatly improved the paper.

### References

- Allais, Lucy (2010) 'Kant's argument for transcendental idealism in the Transcendental Aesthetic'. *Proceedings of the Aristotelian Society*, 110, 47–75.
- Balanovskiy, Valentin (2018) 'What Is Kant's transcendental reflection?'. *Proceedings of the XXIII World Congress of Philosophy*, 75, 17–27.
- Bennett, Jonathan (1966) *Kant's Analytic*. Cambridge: Cambridge University Press.
- Bird, Graham (2006) *The Revolutionary Kant: A Commentary on the 'Critique of Pure Reason'*. Chicago and LaSalle: Open Court.
- Bonjour, Laurence (1998) *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.
- Descartes, René (1964–76) *Oeuvres de Descartes*. Ed. Charles Adam and Paul Tannery. Revised edition. Paris: Vrin/CNRS.
- De Boer, Karin (2020) *Kant's Reform of Metaphysics: The 'Critique of Reason' Reconsidered*. Cambridge: Cambridge University Press.
- Forster, Michael N. (2008) *Kant and Skepticism*. Princeton: Princeton University Press.
- Guyer, Paul (1987) *Kant and the Claims of Knowledge*. Cambridge: Cambridge University Press.

- Hutton, James (2019) ‘Epistemic normativity in Kant’s “Second Analogy”’. *European Journal of Philosophy*, 27, 593–609.
- Kant, Immanuel (1783/1997) *Prolegomena to Any Future Metaphysics*. Trans. Gary Hatfield. Cambridge: Cambridge University Press.
- (1785/2019) *Groundwork for the Metaphysics of Morals*. Trans. Christopher Bennett, Joe Saunders and Robert Stern. Oxford: Oxford University Press.
- (1787/1998) *Critique of Pure Reason*. Trans. and ed. Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.
- (1999) *Correspondence*. Trans. and ed. Arnulf Zweig. Cambridge: Cambridge University Press.
- Kitcher, Patricia (1994) *Kant’s Transcendental Psychology*. Oxford: Oxford University Press.
- McWherter, Dustin (2016) ‘The Moore-Walker dilemma: A critique of transcendental idealism’. *The Philosophical Forum*, 47, 195–206.
- Marshall, Colin (2014) ‘Does Kant demand explanations for *all* synthetic a priori claims?’ *Journal of the History of Philosophy*, 52, 549–76.
- Merritt, Melissa McBay (2015) ‘Varieties of reflection in Kant’s Logic’. *British Journal for the History of Philosophy*, 23, 478–501.
- Millican, Peter (2017) ‘Hume’s fork, and his theory of relations’. *Philosophy and Phenomenological Research*, XCV, 3–65.
- Moore, G. E. (1903–4) ‘Kant’s idealism’. *Proceedings of the Aristotelian Society*, 4, 127–40.
- (1953) *Some Main Problems of Philosophy*. London: George Allen & Unwin.
- Nietzsche, Friedrich (1886/1966) *Beyond Good and Evil*. Trans. Walter Kaufmann. New York: Random House.

- O'Shea, James R. (2019) 'On Sellars' exam question trilemma: Are Kant's premises analytic, or synthetic a priori, or a posteriori?' *British Journal for the History of Philosophy*, 27, 402–21.
- Pereboom, Derk (1990), 'Kant on justification in transcendental philosophy'. *Synthese*, 85, 25–54.
- Pollok, Konstantin (2017) *Kant's Theory of Normativity: Exploring the Space of Reason*. Cambridge: Cambridge University Press.
- Rosenberg, Jay (2005) *Accessing Kant: A Relaxed Introduction to the Critique of Pure Reason*, Oxford: Oxford University Press.
- Russell, Bertrand (1961) *History of Western Philosophy*. 2nd edition. London: George Allen & Unwin.
- Schafer, Karl (2021) 'Transcendental philosophy as capacities-first philosophy'. *Philosophy and Phenomenological Research*, 103, 661–86.
- Smit, Houston (1999) 'The role of reflection in Kant's *Critique of Pure Reason*'. *Pacific Philosophical Quarterly*, 80, 203–23.
- Stern, Robert (2009) *Hegelian Metaphysics*. Oxford: Oxford University Press.
- Tolley, Clinton (2006) 'Kant on the nature of logical laws'. *Philosophical Topics*, 34, 371–407.
- Van Cleve, James (1999) *Problems from Kant*. Oxford: Oxford University Press.
- Westphal, Kenneth R. (2004) *Kant's Transcendental Proof of Realism*. Cambridge: Cambridge University Press.