

ARTICLE

Revisiting the Proof-Structure of Kant's Transcendental Deduction

Hyoung Sung Kim

Stanford University, Stanford, CA, USA
Email: hyongsung@gmail.com

Abstract

There is no consensus concerning how to understand the 'two-step proof structure' (§§15–20, 21–7) of the Transcendental Deduction in the B-edition of the *Critique of Pure Reason*. This disagreement invites a closer examination of what Kant might have meant by a 'transcendental deduction'. I argue that the transcendental deduction consists of three tasks that parallel Kant's broader project of a 'critique' of pure reason; first, an *origin* task to justify reason's authority to use them; second, an *analytical* task that determines the conditions under which this authority *can* be legitimately exercised; and third, a *dialectical* task to determine the conditions under which this authority *cannot* be legitimately exercised. So long as we continue to read the B-Deduction solely in terms of its two-step proof structure, we overlook how Kant's notion of 'critique' constitutes the real grounds for his argumentative strategy there.

Keywords: Kant; transcendental deduction; B-Deduction; two-step proof structure; critique of pure reason

I. Introduction

Kant's B-edition of the Transcendental Deduction of the Categories (the 'B-Deduction') is divided into two steps or parts. The main proof of the first step (§§15–20) is found in §20, while the main proof of the second step (§§21–7) is found in §26. Call these two proofs the 'first proof' and 'second proof' respectively.

One of Kant's main aims for the B-Deduction is to demonstrate the 'objective validity' of the categories, or how the categories as *a priori* concepts of the understanding must apply to empirical intuitions given to us through a distinct, receptive faculty of sensibility (A89–90/B122).¹ In a canonical 1969 paper, Dieter Henrich argued that any adequate interpretation of the B-Deduction must explain its 'two-step proof structure' or why the first proof is necessary but not sufficient to show objective validity and how the second step does so by drawing upon the first (Henrich 1969: 642).² The secondary literature generated by Henrich's interpretative criterion can be organized as follows.

According to *insufficiency* readings, the first proof does not address objective validity, which is demonstrated in the second proof. Henry Allison defends this reading. For Allison, the first proof demonstrates the categories as necessary conditions for

© The Author(s), 2022. Published by Cambridge University Press on behalf of Kantian Review. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

thinking of an object of intuition in general. The second proof then demonstrates objective validity, i.e. the categories as conditions for *cognizing* empirical intuitions.³

According to *aspect* readings, both proofs demonstrate aspects of objective validity. Anil Gomes defends this reading. For Gomes, the first proof demonstrates that we must apply the categories for cognition (validity), while the second proof demonstrates that the categories do actually apply to *objects* of experience (objectivity).⁴

According to *sufficiency* readings, the first proof demonstrates objective validity for an undetermined range of intuitions in general, while the second proof *restricts* or delimits the scope of such validity to empirical intuitions structured by our spatio-temporal forms. Only then, someone like Robert Pippin argues, can Kant establish his real aim for the Deduction, namely, to show that the categories cannot apply to things-in-themselves.⁵

This article aims to argue for a reading of Kant's proof-strategy of the B-Deduction that reconciles the core insights of these three readings. In section 2, I discuss the two-step proof structure. In sections 3–5, I discuss the virtues and limitations of Allison's, Gomes' and Pippin's respective versions of these readings, treating them as representative. Any adequate reading of the B-Deduction must, I argue, integrate their respective core insights.

In section 6, I make headway on such a reading by getting clearer on Kant's argumentative strategy in a 'transcendental deduction'. By revisiting Kant's juridical metaphor of a legal deduction and the transcendental deductions of space and time as well as the transcendental ideas, I argue that the transcendental deduction of the categories consists of three tasks: first, an 'origin' task to justify reason's *authority* to use them by tracing their *a priori* origins to the logical forms of judgement; second, an 'analytical' task that determines the conditions under which this authority *can* be legitimately exercised, namely, that despite their pure origins they are nevertheless limited to appearances as objects of possible experience; and third, a 'dialectical' task to determine the conditions under which this authority *cannot* be legitimately exercised, namely their relation to things-in-themselves (again, despite their pure origins). These tasks parallel Kant's broader project of a 'critique' of pure reason to determine the 'boundaries' of categorical application. So long as we read the B-Deduction solely in terms of a 'two-proof' structure, we overlook how Kant's notion of 'critique' constitutes the real grounds for his argumentative strategy there.

In sections 7 and 8, I sketch my reading of the two-step proof structure. On my view, the first step addresses the origin task (§§15–20), which is the focus of section 7, while the second step addresses the analytical (§§24 and 26–7) and dialectical (§§21–3 and 25) tasks, which is the focus of section 8. I conclude (section 9) by explaining how my reading capitalizes on the respective insights of the three readings in sections 3–5, as well as discussing a tendency to frame the problem of the B-Deduction as foreclosing the possibility of 'cognitive dissonance' between objects given to us through sensibility and such objects as thought by the understanding. If my reading is right, such a tendency misses the mark.

2. The two-step proof structure

In his 1969 paper, Henrich emphasized the significance of claims in §21, titled 'Remark' (or 'Observation'):

In the above proposition [a summary of the result of §20] . . . the beginning of a deduction of the pure concepts of the understanding has been made, in which, since the categories arise **independently from sensibility** merely in the understanding, I must abstract from the way in which the manifold for an empirical intuition is given, in order to attend only to the unity that is added to the intuition through the understanding by means of the category. In the sequel (§26) it will be shown from the way in which the empirical intuition is given in sensibility that its unity can be none other than the one the category prescribes to the manifold of a given intuition in general according to the preceding §20; thus by the explanation of its *a priori* validity in regard to all objects of our senses the aim of the deduction will first be fully attained. (B144–5, my underlining)

In the first underlined sentence, Kant refers to the first proof as only a ‘beginning’ of the deduction that will be ‘fully attained’ in the ‘sequel’, i.e. the second proof. In the second underlined sentence, Kant claims this second proof will complete the B-Deduction by drawing upon the first. The two-step proof structure follows.

The secondary literature has raised two further related constraints. The first is a ‘triviality puzzle’ (Vinci 2014: 179–203). Since Kant’s conclusion in §20 is that ‘all sensible intuitions must stand under categories’ while in §26 his conclusion seems to be the narrower claim that all *human* (spatiotemporally structured) sensible intuitions must so stand, the puzzle is why Kant did not infer the latter directly from the former and included a second step (§§21–7) (see Allison 2015: 328).

Some of Kant’s remarks complicate the triviality puzzle. Given the finality of the conclusion of the first proof, namely, that ‘thus the manifold in a given intuition also necessarily stands under categories’ (B143), readers might be forgiven for thinking that this first proof is the or a transcendental deduction. After Kant denies this initial impression in §21, at the beginning of §26 he seems to then confirm this initial impression by referring to §20 and §21 as ‘the’ transcendental deduction (B159).

A second constraint is suggested by Kant’s claim in §21 that he had to ‘abstract from the way in which the manifold for an empirical intuition is given’ in the first proof: while the first proof abstracts from the specific properties of space and time as determined in the Transcendental Aesthetic, the second proof appeals to these properties for its conclusion. Focusing on space, in the Metaphysical Exposition Kant argues that space is an intuitive unity, or a whole prior to its parts, and thus an *a priori* intuition (not a concept) that originates from pure sensibility (A25/B39–40). In the Transcendental Exposition Kant appeals to this property to explain the possibility of geometry as a synthetic *a priori* science (B41). The second constraint is thus to explain why and how Kant appeals to the doctrines of the Aesthetic in the second proof but not the first.

3. Insufficiency readings

Since for Allison the first proof demonstrates the categories as the necessary conditions for thinking objects of sensible intuitions in general, the problem left open is the possibility of a lack of conformity or ‘cognitive dissonance’ between objects thought by the categories and objects given to us through space and time as forms of our

sensibility (Allison 2015: 9). For Allison, the second proof forecloses this possibility by introducing our sensible forms of space and time to show that the categories are also necessary conditions for *cognizing* empirical intuitions (pp. 69, 436).

Allison's reading has many virtues, of which I emphasize two. First, Allison is sensitive to Kant's important distinction between thought and cognition for our understanding of the B-Deduction (see B146, 148 and 151). Second and relatedly, Allison is sensitive to Kant's emphasis on the importance of *a priori* sensibility as an independent condition for cognition. In general, insufficiency readers emphasize on textual and independent philosophical grounds a distinction between a representation and representing the representation as a representation (see A78/B103, A99 and A120). So: sensibility gives a mere manifold, while the understanding represents this manifold as a manifold. Or, while objects are *given* to us through sensibility, the understanding represents a given object as an object (Allais 2009).

Allison applies this interpretative strategy to explain the second proof. Following Onof and Schulting, Allison identifies sensibility as a source of 'unicity' or a property of space whereby any part of space is immediately and indefinitely intuitable as a part of another, larger space (Onof and Schulting 2015). Allison casts Kant's distinction between 'form of intuition' and 'formal intuition' in the second proof (B160n.) in terms of the unicity of space (qua form of intuition) which originates from sensibility and the representation of this unicity as a unity (qua formal intuition) which is an act of the understanding in the guise of the unity of apperception. Since the categories are conditions for representing all unity, and this unicity through which all empirical intuitions can be given to us can itself be represented by the understanding (in the guise of the productive imagination), the categories must thus also apply to empirical intuitions, i.e. are objectively valid.

A major limitation of Allison's reading is his deflationary reading of the first step. For Allison, the reader must engage in a suspension of disbelief when reading Kant's references to 'cognition', 'intuition', 'object' and 'objective validity' throughout the first step. For Allison, Kant intends these notions in a 'thin' sense, e.g. 'cognition' as referring to thinking an object of a manifold of intuition in general; or by an 'intuition' Kant is referring to a 'manifold of intuition in general' or an intuition considered independently of our sensible forms of space and time as a possible object of thought (Allison 2015: 353).

While Allison's deflationary reading plausibly applies to Kant's remarks concerning the 'I think' in §16, Kant begins §17 by citing the doctrines of space and time as 'formal conditions' of sensibility as shown in the Transcendental Aesthetic, which would suggest that he is transitioning from the unity of apperception as conditions of thinking in §16 to such unity as the conditions for cognition in §17. Kant does go on to identify the understanding as 'generally speaking, the faculty of **cognitions**' (emphasis Kant's) and then claims that 'the unity of consciousness is that which alone constitutes the relation of representations to an object, thus their *objective validity*, and consequently is that which makes them into *cognitions*' (B137, my emphasis). Kant then points out that 'the synthetic unity of consciousness is . . . an objective condition of all cognition . . . something under which *every intuition* must stand **in order to become an object for me**' (B138, my italics).

Consider also the wide-ranging generality suggested by the title of §20: '*all* sensible intuitions stand under the categories, as conditions under which alone their manifold

can come together in one consciousness' (§20; B143, my emphasis), whereby 'all sensible intuitions' include both pure and empirical intuitions. The polysyllogism in §20 also concludes: 'Thus the manifold in a *given* intuition also necessarily stands under categories' (B143, my italics). The cost of the kind of suspension of disbelief Allison is asking of us, on my view, is too steep.

In sum, then, insufficiency readings satisfactorily explain the second proof. They fare less well with the first step and the first proof. Such readings suggest we adopt a suspension of disbelief when reading the first step. But as we shall see, the viability of aspect and sufficiency readings suggests that such a suspension of disbelief is by no means forced upon us.

4. Aspect readings

According to Gomes, the first proof only demonstrates validity or that we must apply the categories for empirical objects (cognition). However, that we must apply the categories need not imply that the categories must so apply (Gomes 2010: 129). For Gomes, the second step then shows that the categories must apply by explaining the unities of space and time in terms of the understanding as producing an 'effect' or determination of sensibility. Because of this shared origin, Gomes argues, 'appearances which are given in space and time thereby fall under the same synthetic unity of apperception which is responsible for synthesising the manifold of intuition in accordance with the categories' (p. 132). Since the unity of space is realized by the unity of apperception, objects must be so constituted such that the categories must apply – because, in short, they cannot be given in a way that escapes such application.

Aspect readers argue that the real lesson of the Deduction is to show how the 'givenness' of intuitions is already determinable, so to speak, for categorical application (see Longuenesse 1998: 214–26). The notion of a 'formal intuition' and the unity of space as a 'determination' or 'effect' of the understanding upon sensibility (see §24) in the second proof makes explicit what was left implicit in the Aesthetic as a science of pure sensibility. As is well-known, the second sentence of B160n. supports this notion of a 'pre-categorical' or 'pre-discursive' application of the unity of apperception when Kant says he 'ascribed this unity merely to sensibility . . . though to be sure it presupposes a synthesis'. In the *Opus Postumum* Kant seems to state this position more clearly, e.g. 'space and time are products (but primitive products) of our own imagination, hence self-created intuitions' (*OP*, 22: 37; see also *Anth*, 7: 167); or how space and time have 'merely a form, a form which we ourselves must make' and that their 'representation is an act of the subject itself and a product of its imagination' (*OP*, 22: 77).

Again, let me highlight two main virtues of Gomes' aspect reading. The first is textual. Unlike insufficiency readings, aspect readings preserve a straightforward reading of the first step. Unlike insufficiency readings, they explain the natural initial impression that the first proof seems to accomplish a deduction of some kind – namely, justifying reason's warrant to use the categories.

Second, aspect readings like that of Gomes frame the B-Deduction as addressing an unacceptable 'subjectivism' or the worry that the objective application of the categories could be explained in terms of contingent facts about our subjective constitution (see also Shaddock 2014, 2015). There is considerable textual evidence to suggest that

Kant had this subjectivist worry in mind while drafting the B-Deduction. The first is found in §27, where he is responding to Schultz. By Kant's lights, Schultz is someone who, while granting the first proof, then argues that the necessary subsumption of appearances under the understanding can be explained in terms of pre-established harmony. The second is less explicit but is found in a canonical passage widely regarded as Kant's formulation of the problem that the Deduction is meant to solve in §13:

[B]ut that [objects of sensible intuition] must also accord with the conditions that the understanding requires for the synthetic unity of thinking is a conclusion that is not so easily seen. For appearances could (*könnten*) after all be so constituted that the understanding would not find them in accord with the conditions of its unity . . . Appearances would nevertheless offer objects to our intuition, for intuition by no means requires the functions of thinking. (A89–90/B122–3)

While there is disagreement concerning the modal status of the possibility Kant is raising, for present purposes I show that Kant has Hume in mind here.⁶ First, throughout the above passage and in the paragraphs that bookend it, Kant uses causality as his example to illustrate the problem. Second, according to the Guyer/Wood translation of the above passage, in Kant's copy of the first edition, Kant appended to the sentence that ends with 'not so easily seen' in the above spectre passage the remark that 'A subjective necessity, habit, would make it worse. An implanted necessity would not prove necessity' (23: 26), and this reference to 'habit' suggests again that Kant has in mind Hume's explanation of categorical necessity in terms of the strength of repeated associations or 'custom' (see also B5 and B20 for the explicit attribution of an explanation in terms of 'habit' to Hume). Third, and relatedly, the insufficiency of custom to justify categorical application is reinforced in the paragraph immediately following the above passage, where Kant tells us one could evade this investigation 'by saying that experience constantly offers examples of a regularity of appearances that give sufficient occasion for abstracting the concept of cause from them, and thereby at the same time thought to confirm the objective validity of such a concept' (A91/B123–4), which, again, is a clear reference to Hume's appeal to custom.

For Kant, Hume and Schultz ultimately explain how the categories must apply by appealing to our subjective constitution – in Hume's case, through our empirical conditions, or in Schultz's case, through how God has made us – which would leave us in a state where 'I would not be able to say that the effect is combined with the cause in the object (i.e., necessarily), but only that I am so constituted that I cannot think of this representation otherwise than as so connected' (B168). For aspect readings, only the second step addresses this worry.

The main worry, then, with aspect readings is the appeal to the unity of apperception to explain the unity of space. Kant seems to take back his endorsement of this strategy in the final sentence of B160 when he emphasizes that the 'unity of this *a priori* intuition belongs to space and time, and not to the concept of the understanding' (B160n.). Aspect readers will argue that by the 'concept of the understanding' Kant is claiming that the unity in question pertains to space and time and not to the *categories*, which leaves open the unity of space as a product of a pre-categorical application of the unity of apperception.

The text is not unambiguous on this issue. For example, in the *Prolegomena* Kant tells us that a 'formal intuition' is 'the essential property of our sensibility by means of which alone objects are given to us' (P, 4: 288; see also A429/B457). Supposing that we grant that such an interpretative move has textual support, there are good grounds to resist them as un-Kantian in spirit. Kant emphasizes that the doctrine of pure sensibility as a source of *a priori* forms independent of the understanding is what makes his philosophy revolutionary within the history of Western philosophy (see P, 4: 376n.). While aspect readings emphasize, rightly, how understanding determines sensibility, the view that the unity of space originates from the understanding does not track Kant's aim in the second step to *restrict* the activities of the understanding by appealing to our sensible forms (A147/B187). If aspect readings are right, it would suggest that, contrary to Kant's intentions, the understanding constrains sensibility and not the reverse. On this issue, insufficiency readings explain better the second constraint, namely, how the Aesthetic supplies an independent constraint in the second proof (see McLear 2015). Without the doctrinal independence of the Aesthetic, we do not have Kant but Hegel or some other German idealist.

Finally, there is textual evidence that Kant would resist this interpretative move:

Despite their dissimilarity, understanding and sensibility by themselves form a close union for bringing about our cognition, as if one had its origin in the other, or both originated from a common origin; but this cannot be, or at least we cannot conceive how dissimilar things could sprout forth from one and the same root. (*Anth*, 7: 177; see also 'On a Discovery', 8: 249–50)

While aspect readers like Gomes suggest that 'it is the nature of the understanding which explains both why we must apply the categories and why the categories must apply' (Gomes 2014: 12), Kant instead seems to argue for a kind of quietism, or that there is no answer to this question.

In addition, while Kant rejects pre-established harmony to explain why the categories must apply to appearances, he suggests his endorsement of such pre-established harmony to explain the unity of our faculties (Letter to Herz, 11: 52; 'On a Discovery', 8: 250). In his letter to Markus Herz in 1789, after indicating his endorsement of such a position, Kant outlines the real task of the Deduction, namely, to show that 'once they are given [that is, our *a priori* forms], however, we are fully able to explain their power of making *a priori* judgments (that is, to answer the question, *quid juris*)' (11: 52). That is, given (i) how space and time are forms of our intuition and their characteristic properties (whole-part priority, unity, singularity) as shown in the Aesthetic; and (ii) how the categories are forms of our thinking and their characteristic properties (necessity, universality) as shown in the Analytic; and (iii) given further the 'fact' that these forms must be able to cooperate to produce synthetic *a priori* cognitions in mathematics and general natural science (see the Introduction) the '*quid juris*' is to show how (i) and (ii) together explain (iii), and what this means for the scope of legitimate categorical application. The problem, then, is not so much the question of the possibility of cognitive cooperation but rather being able to comprehend how these forms together bring about empirical cognition, i.e. *experience* as suggested by our best scientific theories.⁷ I shall return to the significance of this point in the conclusion.

In sum, aspect readings rightly identify ‘subjectivism’ as a threat Kant took seriously in the B-Deduction. However, aspect readings do well with the first proof but not the second. The challenge is to read the second proof while preserving the doctrinal independence of the Aesthetic.

5. Sufficiency readings

According to Pippin, while the first proof establishes objective validity for an indeterminate scope of intuitions in general, the second proof appeals to the Aesthetic to show that, since space and time apply solely to appearances, the scope of categorical application for cognition is thus restricted to such forms and all that is given to us through them (empirical intuitions). As Pippin puts it, while the first step shows that ‘what we do sensibly experience is subject to the categories’, the second step shows that ‘we could not have any knowledge not categorically prescribed’. Only then, for Pippin, ‘the Dialectic can rest, as it everywhere does, on the conclusions of the Deduction (Pippin 1982: 182–3).

Sufficiency readings read the second proof as restricting the scope of the accomplishment of the first proof, as suggested by Kant’s uses of ‘*insofar*’ in the first proof. The first step shows that the categories *only* have significance or can relate to objects when applied to sensible intuitions in general. The second step, by appealing to the Aesthetic, shows that the categories have such significance *only* when applied to sensible intuitions (appearances), thereby ruling out the possibility of applying the categories to cognize things-in-themselves. While these two conclusions sound similar they are not, because they address two different threats. The first step addresses an ‘emptiness threat’ or how the categories might not apply to objects in general, while the second step addresses and precludes a possible ‘overapplication threat’, namely, what Kant calls a ‘natural propensity’ (A642/B670) of reason for ‘transcendental use’ of the categories beyond experience to things-in-themselves. In his discussion of the transcendental deduction of space in §13, Kant says the *a priori* status of space makes the ‘**concept of space**’ ambiguous by inclining us to use it beyond the conditions of sensible intuition, on which account a transcendental deduction of it was also needed above (A88/B120–1) – a consideration such as would also apply to the categories. For sufficiency readings, the two steps parallel Kant’s claim that ‘sensibility *realizes* the understanding’ – the task of the first step – but ‘at the same time it *restricts* it’ – the task of the second step (A146–7/B185–7, my emphasis).

Again, there are two main virtues of Pippin’s reading. First, unlike the other two readings, its restrictive reading of the second proof highlights Kant’s anti-rationalist ambitions in the B-Deduction. Kant tells us the deduction is difficult because it must settle the possibility of metaphysics as a science, ‘whose final aim in all its preparations is directed properly only to the solution’ of the ‘unavoidable problems’ of ‘God, freedom, and immortality’ (A3/B7). Demonstrating objective validity is a means to this final aim. In the Schematism chapter, Kant summarizes the deduction as addressing an overapplication worry:

After what has been shown in the deduction of the categories, hopefully no one will be in doubt about how to decide the question, whether these pure concepts of the understanding are of merely empirical or also of

transcendental use, i.e., whether, as conditions of a possible experience, they relate *a priori* solely to appearances, or whether, as conditions of the possibility of things in general, they can be extended to objects in themselves (without any restriction to our sensibility). (A139/B178)

Finally, in remarks only included in the B-edition prior to §15 of the B-Deduction, Kant says ‘we are now about to make an attempt to see whether we cannot successfully steer human reason’ between the two cliffs of Humean scepticism and what Kant calls ‘enthusiasm’, which ‘dared to make attempts at cognitions that go far beyond the boundary of all experience’ (B127), as well as ‘assign its determinate boundaries, and still keep open the entire field of its purposive activity (B128). Insofar as Kant frames the B-Deduction as a middle course between scepticism and enthusiasm, I take him to seek to show both how the categories are restricted to appearances, thereby vindicating a restricted form of enthusiasm against a general Humean scepticism, as well as how they cannot apply to things-in-themselves, which, in turn, vindicates a restricted form of Humean scepticism concerning our capacity to use reason to cognize things-in-themselves (see also Hatfield 2003).

Moreover, Kant’s shift in language between the first and second steps also supports insufficiency readings. In the first step, Kant emphasizes how the categories must apply to sensible intuitions, e.g. ‘the synthetic unity of apperception . . . is something under which every intuition must **stand in order to become an object for me**’ (§17, B137); ‘all sensible intuitions stand under the categories’ (§20). However, in the second step Kant says that the ‘categories . . . are only rules for an understanding whose entire capacity consists in thinking’ (§21, B145). Both the title and concluding sentence of §22 read: ‘The categories consequently have no other use for the cognition of things except insofar as these are taken as objects of possible experience’, which Kant says ‘is of the greatest importance, for it determines the boundaries of the use of the pure concepts of the understanding in regard to objects’ (B148). In §27, Kant summarizes the conclusion of the B-Deduction as follows: ‘**no a priori** cognition is possible for us except solely of objects of possible experience’ and ‘this cognition . . . is limited merely to objects of experience’ (B165–6). These formulations all suggest a shift in emphasis from showing that the categories have significance when applied to sensible intuitions to showing that the categories *only* have such significance when applied to empirical intuitions, i.e. how reason cannot use them for things-in-themselves.

Second, while aspect and insufficiency readings tend to focus on §26 and the first half of §24, sufficiency readings can incorporate neglected passages in the second step as part of Kant’s general argumentative strategy. Throughout the second step, Kant entertains and rejects various possible domains of categorical application that his rationalist contemporaries would have entertained, including intellectual intuition (see B145; see also B135 and 138); noumena in a positive sense (see B149–50); and the self (§25). While someone like Allison gives voice to a general tendency to regard these texts as ‘digressions’ (Allison 2015: 372), for insufficiency readers – especially emphasized by Frederick Rauscher – these passages are part of the scope-restricting aims of the second step, e.g. first showing that the categories only apply to a discursive and not intuitive understanding; then ruling out noumena in a positive sense to limit the categories to sensibility in general (second half of §23); then ruling out categorical application to the self (§25). Kant is marking boundaries of theoretical

reason: he is arguing that the categories as *a priori* concepts, despite seemingly applying to these representations, cannot apply to them, in ways that preview his discussion in the Paralogisms and the Dialectic. These negative arguments prepare the reader to appreciate the main claim of the second proof, namely, how the categories *only* apply to appearances given to us through space and time and not things-in-themselves (Rauscher 2014: 407–10).

Now, the main problem with sufficiency readings is that Kant's 'beginning' remarks in §21 do not straightforwardly support a restrictive reading of the second proof. Generally, sufficiency readings fail to address the triviality puzzle. Someone like Rauscher, for example, argues that the two steps address the *necessity* and the *universality* of the categories respectively by appealing to this distinction from the Introduction (B3). But while I concede that the text supports this distinction (see the title of §26), the main problem is that Kant thinks this distinction is a pragmatic one: in certain cases, it might be easier to demonstrate the necessity of a judgement by demonstrating universality and then inferring necessity directly or the other way around (B2–3). This leaves the triviality worry in place, or why the first step is only a beginning.

In sum, sufficiency readings rightly emphasize the importance of Kant's broader aim in the *Critique*, which is to deny metaphysical knowledge concerning things-in-themselves, e.g. God, the soul and the beginning of time. Such readings also explain why Kant structured the second step the way he did. However, they do not explain as well as the other two readings the two-step proof structure, and especially the triviality worry. But while Rauscher's appeal to the necessity–universality distinction is mistaken, I think he is on the right track: sufficiency readings need to explain, by appealing to broader considerations within Kant's philosophy, why he thought he needed to establish how the categories apply to a range of intuitions before demonstrating a restricted version of his claim in a further second step by introducing our human, spatiotemporal forms. In what follows, I argue that this task can be readily accomplished by returning to Kant's broader project of a 'critique' of reason in general, in which he is equally concerned to establish the legitimate *and* illegitimate domains of categorical application.

6. Transcendental deductions as an argument strategy

While these three readings cannot all be correct, each of them seems to rest upon a key Kantian commitment: insufficiency readings get right sensibility as an independent constraint; aspect readings get right how the B-Deduction addresses the threat of subjectivism; and sufficiency readings get right Kant's anti-rationalist ambitions for the Deduction. Each reading, then, is getting *something* right about the B-Deduction.

Compare these readings pairwise. Insufficiency and aspect readings explain the 'two-proof' structure of the B-Deduction, i.e. the relation between the proofs in §20 and §26 insofar as the latter does not follow directly from the former. Sufficiency readings, however, explain the 'two-step' structure of the proof. Aspect and sufficiency readings preserve a straightforward reading of the text of the first step. Finally, insufficiency and sufficiency readings explain the second step and especially the second proof in a way that explains how Kant appeals to pure sensibility to

restrict or constrain the *a priori* activities of the understanding. An adequate interpretation of the B-Deduction would reconcile these insights.

In the history of philosophy, it is common for our best thinkers to put forth new philosophical claims in virtue of what they take to be new ways of doing philosophy. Kant is no different. Insofar as Kant reminds us of the novelty, difficulty (P, 4: 260) and ‘depth’ (Axvi) of the B-Deduction and how formulating the argument of a transcendental deduction ‘cost him the most effort’ (Axvi), I believe he is not only referring us to his claims (which are new and difficult), but also, in order to address what he regarded as an equally new kind of philosophical question, namely, the possibility of synthetic *a priori* cognitions, to how he self-consciously develops a transcendental deduction as a completely new kind of philosophical argument within the context of early modern faculty psychology.

The above interpretative impasse is an occasion, I believe, to take a step back and get clearer on Kant’s notion of a ‘transcendental deduction’ as a new method or way of deriving philosophical claims. Perhaps, the thought goes, by a ‘transcendental deduction’ Kant has in mind an argument strategy that consists of three threads or parts; each of these readings, then, would or could be tracking these three parts. Perhaps, continuing this thought, the two-step proof structure is itself derivative of a broader kind of method that rests on considerations internal to Kant’s philosophy.

In what follows, I confirm these thoughts by getting clearer on the relation between Kant’s notion of a transcendental deduction, his project of a ‘critique’ of pure reason and his development of his new science of ‘transcendental logic’.

To do so, I examine Kant’s use of the juridical metaphor of a legal deduction (Henrich 1989); and his remarks concerning the transcendental deductions of space as well as the transcendental ideas. I show that a transcendental-deductive argument strategy involves three tasks: an ‘origin’ task which justifies reason’s authority to use the categories; as well as ‘analytical’ and ‘dialectical’ tasks of showing the conditions under which this authority can be legitimately as well as illegitimately exercised.

In §13, Kant introduces the transcendental deduction in juridical terms as follows:

Jurists, when they speak of entitlements (*Befugnissen*) and claims, distinguish in a legal matter between the questions about what is lawful (*quid juris*) and that which concerns the fact (*Tatsache*) (*quid facti*), and since they demand proof of both, they call the first, that which is to establish the entitlement or the legal claim, the **deduction**. (A84/B116)

In the Guyer/Wood translation of this passage, *Befugnissen* is translated as ‘entitlements’. A better translation, I believe, is ‘authority’, ‘authorisations’, ‘warrants’ or perhaps even, in a legal sense, ‘power’.⁸ Consider how someone might be granted temporary expansive ‘authority’ in emergency circumstances, or how legal challenges could be raised against someone’s having too broad ‘powers’. In Kant’s discussions that demonstrate his knowledge of deduction-writings, he often uses it in the context of justifying authority, e.g. a deduction of an ‘authorised’ publisher (UPB [1785], 8: 79–87), or the ‘authority’ of a sovereign to command subjects to war (MM, 6: 345).

Unlike ‘entitlements’, it is more natural to think of ‘authority’ and ‘power’ as being narrowly or broadly defined in scope, as well as having legitimate and illegitimate exercises (see A148/B187). Kant famously speaks of his critique as instituting a ‘court

of justice' that would 'secure its rightful claims while dismissing all groundless pretensions' (Axi-xii), suggesting reason seeks to determine the conditions under which our theoretical cognitive power can be legitimately applied and illegitimately extended. While Kant uses *Kraft* in contexts where he is speaking of our *cognitive* powers, this double legal and cognitive connotation of 'power' could explain why he decided to appropriate the notion of a deduction-writing to frame what is arguably his most important argument of the *Critique*. These connotations align with Kant's self-characterization of his critical project as determining the boundaries or limits of various cognitive powers and justifying the domain of reason's authority (see e.g. A751-2/B779-80). Kant's tendency to combine legal and cognitive terminology to characterize his important claims suggests he would have encouraged these associations (see A229-30/B281-2; *CPJ*, 5: 174-5).

Henrich has shown that legal deductions during Kant's time were argument forms that justified authority claims by tracing them back to an 'origin' or ground (Henrich 1989). While Henrich's paradigmatic analogy was property rights, in light of my translational point concerning 'authority' a better analogy might be seen in considering as follows a possible 'deduction' of the pardoning power or 'authority' of the President of the United States.

Someone could challenge this power as such: why should the President be able to exercise this power? A 'metaphysical deduction' would justify their authority by tracing it back to the Constitution of the United States as the requisite 'origin' as the 'highest' law of the land. This 'deduction' would thus legitimize, broadly speaking, the President's pardoning authority.

However, the above task would only be, so to speak, a 'beginning' of a deduction insofar as two further challenges could be raised. Granting such presidential authority, the first is: what are the conditions of lawful exercise? We might, in Kantian terms, call this an 'analytical' task, which would restrict, limit or constrain the scope of presidential pardon power, e.g. for federal and not state crimes; and the recipient must agree to accept the pardon.

Completing the analytical task, however, would not be sufficient to determine what Kant might call the 'boundary' of presidential pardoning power. It by no means excludes cases of *unlawful* exercise. One could imagine – without too much difficulty, I should add – a rogue President who would or could believe their pardoning authority extends to the possibility of a self-pardon or 'blanket' pardons for all present and even future crimes by them or their associates. However, the legality of these acts is disputed. Addressing and ruling out these possible uses would be a 'dialectical' task which would entertain certain natural or reasonable conditions where the President might illegitimately seek to *extend* their pardoning power, and then show why such an extension is unwarranted – which would, again, justify these claims by appealing to the Constitution as the highest law of the land. While both analytical and dialectical tasks separately identify 'limits' (*Schranken*) of presidential authority, only when these two tasks are accomplished together would our analogue-Kant think he can firmly draw a 'boundary' (*Grenze*) that presidential pardoning authority cannot exceed (see Howard 2022).

This reading of a 'transcendental deduction' argument strategy in terms of these three tasks can be corroborated by turning to Kant's deductions of space and the transcendental ideas. In §13 Kant tells us that *all a priori* representations, which include

the categories but also space and time as well as the transcendental ideas, require deductions (A85–6/B118; for the ideas, see A669/B697). He also says he has already accomplished deductions of space and time (A88–9/B120–1). Even in the case of the ideas, while Kant concedes that they lack the kind of objective validity possessed by the categories, a deduction is still required to demonstrate their ‘authorised’ regulative function with respect to the unity of the understanding, or what he calls their ‘indeterminate’ objective validity (A671/B699).

In the cases of space and time and the transcendental ideas, however, accomplishing their origin tasks is comparatively easier than doing so for the categories. Focusing on space, for Kant its *a priori* origins are shown through a metaphysical exposition of sensibility, which is then confirmed by geometry as a (synthetic *a priori*) science of space; geometry, in turn, ‘nevertheless follows its secure course through strictly *a priori* cognitions without having to beg philosophy for any certification of the pure and lawful pedigree of its fundamental concept of space’ (A87/B120). Thus, in the case of sensibility the expositions themselves imply their transcendental deduction:

We therefore assert the **empirical reality** of space (with respect to all possible outer experience), though to be sure its **transcendental ideality**, i.e., that it is nothing as soon as we leave aside the condition of the possibility of all experience, and take it as something that grounds the things in themselves. (A28/B44)⁹

Thus, the authorized use of space for appearances is demonstrated by geometry and its exact application to nature (empirical intuitions); the unauthorized use of space for things-in-themselves is demonstrated by considerations that suggest mathematics is not analytic but *synthetic a priori*. As Kant says elsewhere, ‘the principles of the transcendental aesthetic . . . [show that] space and time are the conditions of the possibility of all things as appearances’ (the analytical task) ‘as well as [showing] the restriction of these principles, namely that they cannot be related to things in themselves’ (the dialectical task) (A149/B189, my emphasis). Since the authorized use of space is for pure and empirical intuitions, it cannot be used for objects beyond experience, e.g. the soul or God, as some of Kant’s contemporaries (e.g. Crusius) would have believed (see Messina 2015).

For the transcendental ideas, while Kant thinks their *a priori* origin is established by a metaphysical deduction (see A321/B378), such an origin is already demonstrated by what he calls our ‘natural predisposition’ for metaphysics. Again, as Kant does for space, our use of the ideas is justified by metaphysics as a kind of ‘reality’: ‘metaphysics is real (*wirklich*), if not as a science yet as a *natural predisposition* (*metaphysica naturalis*)’ (B21, my italics), which for Kant validates the claim that the ideas must have some authorized use for reason – namely, within practical reason (see *P*, 4: 362–3). While beyond the scope of this article, Kant seems to think this predisposition is partly a psychological fact about all human beings as well as an *a priori* feature of our rationality in general. In the Appendix Kant then accomplishes his analytical task of showing how the ideas only have a regulative use for systematic inquiry into nature (A672–88/B700–16), as well as the dialectical task of showing how, when used incorrectly (as constitutive of experience), they lead to errors of ‘lazy reason’, ‘perverted reason’ and an unjustified ‘anthropomorphism’ (A689–703/B717–30).

Before turning to the text, let me now explain these tasks in terms of the transcendental deduction of the categories. Kant takes Hume to raise a challenge to reason's *a priori* presumption of authority to apply concepts such as causality and substance (see A760/B788). A transcendental deduction would thus first address this origin task by adducing an *a priori* source – the original-synthetic unity of apperception – as an act of 'pure spontaneity'.

Accomplishing this origin task would still only be a 'beginning' of a deduction, however, for such an origin does not yet entail answers to the 'analytical' and 'dialectical' tasks. The analytical task would demonstrate the conditions of lawful or legitimate exercises of this authority, namely, in relation to empirical objects given to us through sensibility, and thus corresponds to the general task of the Transcendental Analytic, which shows 'the understanding can never accomplish *a priori* anything more than to anticipate the form of a possible experience in general, and, since that which is not appearance cannot be an object of experience, it can never overstep the limits of sensibility, within which alone objects are given to us' (A246–7/B303).

Second, however, given reason's 'natural predisposition' for the unconditioned, transcendental use of the categories, a deduction would also need to address the dialectical task, which entertains and explains away certain natural conditions under which reason seeks to illegitimately extend its authority. This part of a deduction thus corresponds to the part of the transcendental logic that Kant calls a 'discipline' (see A797/B825), which 'serves for the determination of boundaries' and has only the 'silent merit of guarding against errors' (A795/B823) – and thus 'leaves room' open for a way in which the transcendental use of the categories can find its true and purposive realization within practical reason for objects given to us through the moral law.

So long as we continue to understand the relation between the two steps in terms of, say, strict logical dependence, I think we will continue to be puzzled by it. On the view I am urging, however, we should think of the proof-structure of the B-Deduction, the structure of Kant's transcendental logic and the project of a 'critique' of pure reason as one and the same, namely, to show what reason can and cannot cognize through the categories alone.

7. The first step and the origin task

In what remains, I shall corroborate my reading with the text. Namely, I shall sketch a reading of the B-Deduction in terms of addressing the three tasks just distinguished – showing how the first corresponds to Kant's 'metaphysical' deduction and the others to their 'transcendental' deduction.

The origin task is addressed in the first step. Kant begins §15 by characterizing the understanding as a power for combination (B129–30). He claims further that this combination presupposes a higher unity that 'itself contains the ground' of both categorical unity as well as the 'unity of different concepts in judgments, and hence of the possibility of the understanding, even in its logical use' (B130–1) – which, for Kant, is the original-synthetic unity of apperception (§16).

To justify this unity as an act of pure spontaneity, Kant argues as follows. Since one cannot think of a representation unthinkable by any subject – for by the very act by which I think of such a representation, I must think of it – to think of an unthinkable

representation is impossible. Thus, the 'I think' or my thinking must be able to accompany all my representations (B131–2). Kant calls this property of the 'I think' an 'analytic unity' of apperception. It is a condition for all my possible representations – which includes thoughts (concepts, judgements, inferences) as well as intuitions (pure and empirical). In the above passage from §16 Kant then applies this 'I think' claim to all my possible intuitions: 'Thus all manifold of intuition has a necessary relation to the I think in the same subject in which this manifold is to be encountered'. If a subject could not become conscious of herself as common to all the parts of any complex representation – whether that be a pure intuition of a line, the logical unity of an inference, or the different parts of an empirical intuition of a house – she would have no representation at all, just a 'multicolored, diverse self' insofar as she is conscious (B134). Analytic unity is a condition for all my possible representations; an analytic unity of *apperception* is a power to bring about such unities. In §19, Kant then shows that the logical forms of judgement from general logic are just those forms or modes of analytical unity through which we represent an object in general (see A79/B105). The authority of the categories is thus supplied or grounded by the *a priori* status of logic as a science of the form of thinking in general, one that completely abstracts from all sensible conditions (A54–5/B78–9). Moreover, to show this *a priori* origin is, for Kant, to show that these forms must have some objectively valid use, whether it is their possible application to intuitions given to us through sensibility or their applications in practical reason through the moral law.¹⁰

One objection is that what I am describing seems to be what Kant calls his 'metaphysical deduction', which is generally thought to have been accomplished in §10. However, it is also possible to read those passages not as an argument but as anticipating a task that must be accomplished by the transcendental deduction. Moreover, insofar as Kant calls the metaphysical deduction a 'deduction' which addresses a question of authority, my proposal is simply that it is a deduction in the sense of addressing the 'origin' task, or the question of what grounds reason's authority to use the categories.¹¹

Consider next the passage in which Kant introduces the distinction between a metaphysical and transcendental deduction:

In the **metaphysical deduction** the origin of the *a priori* categories in general was established through their complete coincidence with the general logical functions of thinking, in the **transcendental deduction**, however, their possibility as *a priori* cognitions of objects of an intuition in general (§§20, 21) was exhibited. (B159)

One plausible way to read the parenthetical is that Kant is drawing our attention to §20 and §21 as concerning the metaphysical and transcendental deductions respectively. I think this is not so implausible; by 'metaphysical deduction' in §20 Kant could have in mind the passages 'That action of the understanding, however, through which the manifold of given representations (whether they be intuitions or concepts) is brought under an apperception in general, is the logical function of judgments (§19)' and 'the categories are nothing other than these very functions for judging insofar as the manifold of a given intuition is determined with regard to them (§13)'. The reference to §21, in turn, could be to Kant's remark there that the first

proof ‘indicates, therefore, that the empirical consciousness of a given manifold of one (*Einer*) intuition stands under a pure *a priori* self-consciousness, just as empirical intuitions stand under pure sensible one, which likewise holds *a priori*’ (B144). Kant could be read as previewing the second proof in a compressed, preliminary manner, taking the reference to how empirical intuitions stand under space and time as suggesting the synthesis of apprehension, which is the first premise of the second proof. All in all, in any case, the first step justifies reason’s authority to use the categories, and thus addresses the challenge raised by Humean scepticism (Kohl 2018).

8. The second step: the dialectical and analytical tasks

If I am right to read the first step as addressing the origin task, we can see why this would nevertheless only be the ‘beginning’. For the understanding, as Kant says, is in and of itself ‘empty’ where ‘nothing manifold is given’ (B135); a manifold given to it through sensibility is required to ‘realize’ the activities of the understanding and give it significance or meaning, i.e. a relation to objects. Demonstrating authority is not yet to determine the conditions of legitimate and illegitimate exercise of such authority. As Kant tells us, ‘since the categories arise **independently from sensibility** merely in the understanding’ – a reference to the focus on the *a priori* origin in the first proof – he has thus far attended ‘only to the unity that is added to the intuition through the understanding by means of the category’ (B144). Kant thus indicates a shift from demonstrating the authority of the categories by identifying their *a priori*, synthetic origins in the first step (namely, so that they *must* be objectively valid) to analysing the legitimate and illegitimate conditions of exercising this authority or objectively valid use in the second.

Here, Kant seems to get ahead of himself. He informs us that the analytical task will be completed ‘in the sequel (§26)’, telling us that ‘from the way in which the empirical intuition is given in sensibility’ – that is, through our forms of space and time – it will be shown ‘that its unity can be none other than the one the category prescribes to the manifold of a given intuition in general according to the preceding §20; thus by the explanation of its *a priori* validity in regard to all objects of our senses the aim of the deduction will first be fully attained’ (B144–5). Space and time are independent conditions that stipulate the sole domain of authorized use of the categories. For reasons that are not entirely clear, Kant seems to think that the dialectical task must be addressed first.

Perhaps Kant had what would have been his largely rationalist-leaning readers in mind who, having just been told in the first step that the categories are grounded *a priori* completely independently of sensibility, would then expect Kant to vindicate a rationalism whereby the categories could be used to cognize things-in-themselves. That is, the worry to be precluded is a misunderstanding that could be brought about by the determination of the *pure* origins of the categories, for the use of cognizing things-in-themselves.

After his remarks concerning the fact that the first step is only a ‘beginning’ Kant demonstrates the tasks explained earlier in my discussion of sufficiency readings. By emphasizing that the categories do not apply to intellectual intuition (see B145; also B135 and 138), Kant emphasizes how objects given to us are subject to a necessary, separate act of synthesis that proceeds from part to whole. By denying the categories

application to noumena in a positive sense (§23, B149–50), he emphasizes the need to be able to *think* noumena in a negative sense ‘in order not to extend sensible intuition to things in themselves’ and again ‘to limit the objective validity of sensible cognition (for the other things, to which sensibility does not reach, are called noumena just in order to indicate that those cognitions cannot extend their domain to everything that the understanding thinks’ (A254/B310). Kant then finally excludes the possibility of categorical application to the self as thing-in-itself and to the unity of apperception (§25). In these sections, Kant is addressing the dialectical task of not extending the categories beyond appearances.

Having established the boundaries of categorical application, Kant returns to the analytical task in the pivotal §26 and the first half of §24. The main task is to read this second proof as showing how the understanding does not derive the categories from experience, but rather the categories are conditions for the possibility of experience. The second proof begins with a premise concerning the synthesis of apprehension, or an act of combining representations that gives rise to empirical intuitions. Kant then argues this synthesis of apprehension takes place in accordance with space and time as forms of our intuition. The crucial premise is that space and time are not just forms of intuition, but what Kant calls ‘formal intuitions’ or representations ‘with the determination of the **unity** of this manifold [of intuition] in them’ (B160). Since space and time are themselves unities, and the first step has shown that, wherever there is unity of consciousness, the categories must apply, it follows that the categories must apply to the unities of space and time and whatever is given to us through them, i.e. empirical intuitions.

As is well-known, the crucial move of the second proof is Kant’s distinction between ‘forms of intuition’ and ‘formal intuitions’, which, in turn, rests upon a reading of the transcendental synthesis of the imagination or the productive imagination in §24 (B160n.). In my earlier discussion, I urged that we need to read this second proof as limiting cognition to appearances in a way that appeals to space and time as independent forms of constraint. While a detailed treatment of either of these notions as well as their relationship is beyond the scope of this article, for present purposes I shall focus on the productive imagination. I shall explain how Kant’s assumption of mathematics and natural science as the *only* legitimate sources of synthetic *a priori* cognition (call these the ‘established sciences’, see B128) explains the restrictive move of the second proof.¹²

In remarks Kant included only in the Introduction to the B-edition, in a section titled ‘The general problem of pure reason’ (B19), he reiterates that the possibility of metaphysics as a synthetic *a priori* science rests upon the explanation of the possibility of two sciences: pure mathematics (Euclidean geometry) and pure natural science. By ‘pure natural science’ Kant is referring to the transcendental or ‘general’ part of natural science, which includes the principles of pure understanding and in particular the dynamical principles, such as the principles of causality and substance in the Analogies of Experience, as expressing the most fundamental laws of nature, and is treated at length in the Principles of Pure Understanding.

In the Aesthetic, again focusing on space, the Metaphysical Expositions show that space must be an ‘all-encompassing singular space’ or an *a priori*, synthetic, intuitive whole prior to its parts; the Transcendental Exposition then vindicates the Metaphysical Exposition by showing how such a property of space explains geometry

as a legitimate source of synthetic *a priori* cognitions. In the Aesthetic Kant thus explains ‘from below’, as it were, how space and time must be *a priori* intuitions with an intuitive, unified form, in order to explain how pure mathematics and the general doctrine of motion or pure mechanics are sources of synthetic *a priori* cognition.

The second proof addresses ‘from above’, then, the intellectual conditions in which the categories can be legitimately exercised within, and in accordance with, the two established sciences. An explanation of how the categories are used in those sciences would thus provide a complete explanation of the sole, authorized conditions of categorical use within theoretical reason, thus completing Kant’s goal in the Deduction of explaining the possibility of nature as an object of scientific inquiry.¹³ And my proposed interpretative strategy is to explain Kant’s appeal to the ‘transcendental synthesis of the imagination’ as addressing the possibilities of these two sciences together, in terms of what we might call indispensability arguments. In contemporary literature, the canonical formulation of this argument is found in Quine and Putnam, albeit in terms of ontological commitments, i.e. we are ‘ontologically committed’, so to speak, to mathematical notions such as fields, sets, lines, points and so forth, insofar they are indispensable to mathematical explanations that figure in empirical scientific laws.¹⁴ I take Kant to be making a similar move in cognitive terms.

We would be cognitively committed, so to speak, to an account of our faculties that would explain how products of mathematics (as constructions of pure intuition) figure within empirical scientific laws (that apply to empirical intuitions). The synthesis of the ‘productive imagination’ is Kant’s regressive solution to this problem, as a special *a priori* faculty that is part sensible, part intellectual that we must be able to exercise if we are to explain how certain products of the imagination turn out to be *true* or valid of objects in general. The imagination in general is a ‘faculty for representing an object even **without its presence** in intuition’ (B151). Mathematical concepts such as lines and circles, understood solely as constructions through pure intuition, still lack objective validity in the sense that they could be considered *Hirngespinnste* or figments of the imagination. What distinguishes products of an empirical, reproductive imagination (unicorns), from a non-empirical (‘transcendental’) productive imagination, is how the latter, despite being an imaginative capacity, must nevertheless be considered objective in the sense of being *about* objects – which, more concretely, is exercised when our geometric notions are realized within our best empirical-scientific theories.

Kant thus appeals to this productive imagination as the only kind of cognitive faculty that could explain the indispensability of mathematics for empirical science, or how *a priori* constructed notions in the former in pure intuition (e.g. ellipses) constitute solutions in terms of their physical realizations in the latter, or application to empirical intuitions (e.g. orbits of planetary bodies). Absent this faculty, it would be a miracle that pure products of our imagination such as lines, circles and other geometric objects – that are cognized with ‘immediate’ intuitive certainty independent of experience – also just happen to consistently figure within the explanatory resources of our best scientific explanations. Insofar as the human imagination is exercised mathematically (pure intuitions of space and time, mathematical categories, governing our possible *perceptions*) and empirico-scientifically (empirical intuitions, dynamical categories, governing possible *experience*), Kant thinks that it must be no accident that such an exercise yields *true* or objective cognitions, in

the sense of producing the kind of objects ('spontaneously' or as originating from our subjective forms) that also turn out to be necessarily applicable to possible experience.

The only legitimate application of the categories is what is implicated in the exercise of this productive imagination, i.e. in the application of the mathematical categories (geometry) and the dynamical categories (pure natural science) to explain the possibility of nature. The full, complete account of how *all* the categories – mathematical and dynamical – apply solely to objects of empirical intuitions given to us through space and time is implied by the transition from the synthesis of apprehension as conditions for perception (mathematical categories) to synthesis in general as conditions for the possibility of experience (dynamical categories). That is, 'all synthesis, through which even perception itself becomes possible [i.e., the synthesis of apprehension] stands under the categories, and since experience is cognition through connected perceptions, the categories are conditions of the possibility of experience, and are thus also valid *a priori* of all objects of experience' (§26, B161; my underlining), which is a clear reference to the Analogies of Experience, of which the principle as stated in the B-edition is simply: 'Experience is possible only through the representation of a necessary connection of perceptions' (A176/B218). The reference to the Analogies of Experience in this second proof suggests that this second proof is an abstract formulation of a general proof strategy that Kant will pursue more concretely in the Principles of Pure Understanding.

Through this second proof, then, Kant seeks to demonstrate the kind of categorical application necessarily involved in the productive imagination as exercised within the established sciences. The second proof shows how the mathematical categories are required for mathematical construction – the latter of which, in turn, are applicable only to empirical intuitions through the dynamical categories as containing the highest, most general principles of nature. Through this proof, Kant can complete the explanation of the possibility of pure natural science and the possibility of nature in a formal sense (see *P*, 4: 318) – and thereby complete the analytical task of demonstrating that the authorized theoretical use of the categories is limited to the kind of 'experience' that is found within and in accordance with the established sciences.¹⁵

9. Conclusion

I now conclude by comparing my reading with insufficiency, sufficiency and aspect readings, and showing how my reading has implications for the kind of problem that we should read Kant as addressing in the B-Deduction.

First, my reading retains the core insight of insufficiency readings, namely, how the second proof appeals to the doctrines of the Aesthetic as an independent constraint upon the activities of the understanding. But unlike insufficiency readings, mine preserves a straightforward reading of the first step.

Second, my reading retains the core insight of aspect readings, namely, that the first step addresses an aspect of objective validity. In my reading, the first proof shows that reason does have the authority or warrant to use the categories in virtue of their *a priori* origins. But unlike aspect readings, mine rejects the appeal to a common ground between the unity of space and apperception to explain the second step. On my view, Kant's main intention in the second step is to deploy the doctrines of

the Aesthetic to distinguish between the conditioned activities of the understanding, namely being restricted to objects of possible experience (the analytical task), and the unconditioned activities of reason (the dialectical task).

Third, my reading fills a lacuna in sufficiency readings, namely, by providing a framework for understanding the two-step proof structure in terms of Kant's broader project of a critique of pure reason – to explain the boundaries of pure reason – as well as in terms of Kant's transcendental logic and its division into an Analytic and a Dialectic. My reading also urges that a complete understanding of the restrictive move of the second proof requires an essential appeal to Kant's engagement with the mathematical and physical sciences of his time.

I conclude, lastly, with the consequences of my reading for understanding the real problem that Kant seeks to answer in the B-Deduction – an issue I foreshadowed in my discussion of aspect readings in section 4. Traditionally, the Deduction is framed – typically by insufficiency and aspect readers – as occasioned by a problem of Kant's own making, namely, his doctrine of 'discursivity' or of the fact that cognition requires the cooperation of two independent faculties each with their distinct *a priori* forms. Given these two faculties and their distinct forms, it is said, the problem is the possibility of a 'spectre' or the possibility of a lack of conformity between these two faculties: objects could be given to us through space and time in a manner that escapes categorical application. Such a spectre, if genuine, would threaten the possibility of cognition. The Deduction precludes this spectre and affirms that cognition is possible.

I suspect that this way of reading the B-Deduction is attractive if one approaches it for answers concerning, say, the possibility of non-conceptual content (intuitions without concepts). The way the problem is framed leads to problems concerning the ground of the unity of space and its relation to the unity of apperception. However, if my reading is right, then we can and should set aside this way of reading the B-Deduction. For we have already seen in my discussion of aspect readings that Kant seems to endorse a pre-established harmony concerning the ground of the unity of our faculties. Moreover, if cognitive dissonance really were the problem, Kant's appeal to a 'transcendental synthesis of the imagination' as a mediating faculty between sensibility and understanding would be a very 'easy' answer to what is meant to be a difficult problem; we might wonder why *this* faculty is the right kind of faculty to ensure their cooperation.

On my view, the real problem that animates Kant in the B-Deduction – as suggested by the title of the argument as the 'transcendental deduction of the *categories*' – concerns the peculiar status of the *categories* as *a priori* concepts. On the one hand, Kant wants to show how the categories cannot be derived from experience but must instead make experience possible in virtue of their *a priori* status. On the other hand, he wants to show that, despite this *a priori* status, their only legitimate domain of application within theoretical reason is for empirical intuitions and not things-in-themselves. Finally, Kant wants to show that the categories cannot be used to cognize things-in-themselves because he seeks to 'keep open the entire field' of 'reason's purposive activity' (B128) – namely that the true and purposive use of these categories to cognize things-in-themselves is to be found in practical reason.

The real difficulty of the B-Deduction, then, is recognizing that these three tasks must be conjointly accomplished. And these three tasks correspond to what I have

called the origin, analytical and dialectical tasks. And only by accomplishing these tasks together through a ‘transcendental deduction’ can Kant be confident that he has ‘assigned the determinate boundaries’ within which the categories may legitimately and not legitimately be applied – and thereby complete a genuine ‘critique’ of pure reason.¹⁶

Notes

1 All references to the *Critique of Pure Reason* use the standard A/B pagination. All other references to Kant’s writings refer to volume and page numbers of the Academy edition. Translations are from the Cambridge editions of Kant’s works. Abbreviations are the following: *Anth* = *Anthropology from a Pragmatic Point of View*; *CPJ* = *Critique of Power of Judgement*; *MM* = *Metaphysics of Morals*; *P* = *Prolegomena to any Future Metaphysics that will be Able to Come Forward as Science*; *OP* = *Opus Postumum*; *UPB* = *On the Wrongfulness of Unauthorised Publication of Books*.

2 By ‘Transcendental Deduction’ or simply ‘Deduction’ I am referring to the text of the argument in the *Critique*, while by ‘transcendental deduction’ or simply ‘deduction’ I am referring to the general argument strategy.

3 Allison 2015; see also Evans 1990; Onof and Schulting 2015; Allais 2017. Thomas Vinci calls insufficiency readings ‘bifurcation approaches’ insofar as such readings must distinguish on Kant’s behalf two senses of ‘intuition’ that are respectively addressed in the two steps (Vinci 2014: 7–8).

4 Gomes 2010; see also Shaddock 2015; Williams 2018.

5 Pippin 1982; see also Hatfield 2003; Rauscher 2014. Guyer 2010 also seems to endorse a version of this reading insofar as he distinguishes the two steps in terms of an ‘abstract’ and ‘concrete’ proof for objective validity, his doubts concerning the two-step proof structure notwithstanding.

6 Lucy Allais reads the possibility as metaphysical (Allais 2016: 7–9), while Gomes reads this possibility as epistemic (Gomes 2014: 6). On my view, in Kant’s technical terms he has in mind the possibility that the *dynamical* categories may not apply to cognition, i.e. perceptions, without being unified in a whole of possible experience; on this, see Friedman 2014.

7 Here my reading resembles Ameriks’ ‘regressive’ reading of the B-Deduction. See Ameriks 1978.

8 In the Glossary of the Cambridge Edition *Practical Philosophy* volume, *Befugniss* is translated as ‘authorisation, warrant’.

9 Here I am indebted to Merritt 2010.

10 Kohl calls this the ‘apriority-validity’ link: ‘one cannot drive a wedge between the idea that the metaphysical deduction proves the a priori origin of the categories and the idea that the [metaphysical deduction] also proves their validity’; for Kant, ‘since concepts that spring solely from actions of the pure intellect are by definition a priori, the notion of an inherently erroneous, invalid a priori concept is an oxymoron’ (Kohl 2018: 378).

11 Compare with Proops 2003, who argues that the metaphysical deduction is a deduction in a broad sense of a ‘derivation’, e.g. as with a *physiological* derivation of the categories from comparison, abstraction and reflection.

12 Here I am indebted to the work of Michael Friedman (2014).

13 A reader has suggested that this may make it sound as if it were illegitimate to use the categories outside of scientific claims. On Kant’s behalf, I would distinguish between the use of these categories by what he calls the ‘common understanding’ (see B3) and the use of these categories by a critical thinker, i.e. a thinker engaged in the Kantian project of justifying both their pure origins *and* how such pure origins are consistent with their being restricted to objects of possible *experience*. In the context of this latter project, then, I think what Kant wants to say is that, given that mathematics and general natural science are sources of synthetic a priori cognitions, the correct use of the categories for cognizing empirical objects is exhibited in those sciences. Certainly, however, Kant also wants to maintain that true, purposive use of the categories for cognizing things-in-themselves is vindicated in practical reason – something that Kant actually thinks the common understanding can discover (see *P*, 4: 391).

14 For canonical expression, see Putnam 1979; for discussion of Quine’s formulation of the argument, see Colyvan 2001.

15 On the significance of the possibility of nature, or the 'cosmology of experience', for the second step, see Laywine 2020.

16 For helpful comments, discussion and advice I am grateful to R. Lanier Anderson, Richard Aquila (in his capacity as editor of *Kantian Review*), Michael Friedman, Graciela De Pierris, Jessica Williams, Allen Wood and audience members of Michael Friedman's Winter 2022 first *Critique* class. I am especially grateful to the anonymous reviewer for this journal whose comments substantively improved this article.

References

- Allais, L. (2009) 'Kant, Non-Conceptual Content, and the Representation of Space'. *Journal of the History of Philosophy*, 47(3), 383–413.
- (2016) 'Conceptualism and Nonconceptualism in Kant: A Survey of the Recent Debate'. In D. Schulting (ed.), *Kantian Nonconceptualism* (London: Palgrave Macmillan), 1–25.
- (2017) 'Transcendental Idealism and the Transcendental Aesthetic: Reading the Critique of Pure Reason Forward'. In J. R. O'Shea (ed.), *Kant's Critique of Pure Reason: A Critical Guide* (New York: Cambridge University Press), 46–63.
- Allison, H. (2015) *Kant's Transcendental Deduction: An Analytical-Historical Commentary*. New York: Oxford University Press.
- Ameriks, K. (1978) 'Kant's Transcendental Deduction as a Regressive Argument'. *Kant-Studien*, 69(1–4), 273–87.
- Colyvan, M. (2001) 'The Quinean Backdrop'. In *The Indispensability of Mathematics* (New York: Oxford University Press), 21–38.
- Evans, J. C. (1990) 'Two-Steps-in-One-Proof: The Structure of the Transcendental Deduction of the Categories'. *Journal of the History of Philosophy*, 28(4), 553–70.
- Friedman, M. (2014) 'Laws of Nature and Causal Necessity'. *Kant-Studien*, 105(4), 531–53.
- Gomes, A. (2010) 'Is Kant's Transcendental Deduction of the Categories Fit for Purpose?' *Kantian Review*, 15(2), 118–37.
- (2014) 'Kant on Perception, Naive Realism, Non-Conceptualism, and the B-Deduction'. *Philosophical Quarterly*, 64(254), 1–19.
- Guyer, P. (2010) 'The Deduction of the Categories'. In P. Guyer (ed.), *Cambridge Companion to Kant's Critique of Pure Reason* (Cambridge: Cambridge University Press), 118–51.
- Hatfield, G. (2003) 'What were Kant's Aims in the Deduction?' *Philosophical Topics*, 31(1/2), 165–98.
- Henrich, D. (1969) 'The Proof-Structure of Kant's Transcendental Deduction'. *Review of Metaphysics*, 22(4), 640–59.
- (1989) 'Kant's Notion of a Deduction and the Methodological Background to the First Critique'. In Eckart Förster (ed.), *Kant's Transcendental Deductions* (Stanford, CA: Stanford University Press), 29–46.
- Howard, S. (2022) 'Kant on Limits, Boundaries, and the Positive Function of Ideas'. *European Journal of Philosophy*, 30(1), 64–78.
- Kohl, M. (2018) 'The Metaphysical Deduction and the Shadow of Humean Skepticism'. *Kant-Studien*, 109(3), 367–94.
- Laywine, A. (2020) *Kant's Transcendental Deduction: A Cosmology of Experience*. Oxford: Oxford University Press.
- Longuenesse, B. (1998) *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Trans. C. Wolfe. Princeton: Princeton University Press.
- McLear, C. (2015) 'Two Kinds of Unity in the Critique of Pure Reason'. *Journal of the History of Philosophy*, 53(1), 78–110.
- Merritt, M. M. (2010) 'Kant on the Transcendental Deduction of Space and Time: An Essay on the Philosophical Resources of the Transcendental Aesthetic'. *Kantian Review*, 14(2), 1–37.
- Messina, J. (2015) 'Conceptual Analysis and the Essence of Space: Kant's Metaphysical Exposition Revisited'. *Archiv für Geschichte der Philosophie*, 97(4), 416–57.
- Onof, C. and D. Schulting (2015) 'Space as Form of Intuition and as Formal Intuition: On the Note to B160 in Kant's Critique of Pure Reason'. *Philosophical Review*, 124(1), 1–58.
- Pippin, R. (1982) *Kant's Theory of Form*. New Haven and London: Yale University Press.

- Proops, I. (2003) 'Kant's Legal Metaphor and the Nature of a Deduction'. *Journal of the History of Philosophy*, 41(2), 209–29.
- Putnam, H. (1979) 'What is Mathematical Truth'. In *Mathematics, Matter, and Method: Philosophical Papers*. 2nd edn. (Cambridge: Cambridge University Press), 60–78.
- Rauscher, F. (2014) 'The Second Step of the B-Deduction'. *European Journal of Philosophy*, 22(3), 396–419.
- Shaddock, J. (2014) 'Kant and the Most Difficult Thing that Could Ever be Undertaken on Behalf of Metaphysics'. *History of Philosophy Quarterly*, 31(1), 41–58.
- (2015) 'Kant's Transcendental Idealism and his Transcendental Deduction'. *Kantian Review*, 20(2), 265–88.
- Vinci, T. (2014) *Space, Geometry, and Kant's Transcendental Deduction of the Categories*. New York: Oxford University Press.
- Williams, J. (2018) 'Kant on the Original Synthesis of Understanding and Sensibility'. *British Journal for the History of Philosophy*, 26(1), 66–86.