

2

Kant's Deduction *From Apperception*

2.1 Introduction

In my previous book on the Deduction, *Kant's Deduction and Apperception: Explaining the Categories* (Schulting 2012; henceforth in this chapter abbreviated as KDA), I presented a narrowly focused interpretation of the so-called 'first step' of the B-Deduction (running through §§15–20), including the introductory sections 13 and 14. I also dedicated a chapter to §10, i.e. the Third Clue section, that is, the section in the Metaphysical Deduction (MD) where Kant catalogues the categories (A76–83/B102–9), after having furnished the famous guiding thread (*Leitfaden*) to finding the categories. The central hypothesis I advanced was—and no Kantian in his or her right (orthodox) mind, bar Klaus Reich (2001) and a few others following in his wake, has dared to make this claim so far—that, to put it very boldly, in the Transcendental Deduction (TD) Kant effectively derives the categories 'from scratch'.¹ This is not entirely accurate, as Kant of course starts with the premise, or undisputed fact,² of the discursivity of the human mind, the very general characteristics of which he first lists in the first two sections of MD (leading up to the table of judgement), which I did

not discuss in the book (I believe the arguments in the first two sections of MD are exhaustively and conclusively dealt with by Wolff 1995). So to be more precise, contrary to the standard reading, my claim was that in TD Kant derives the categories from the discursivity of the human mind, or, from the capacity to think, thus confirming the validity of the *Leitfaden* provided in MD, which argues that the table of judgement, or more precisely the table of the discursive functions of thought in judgement, and the table of categories neatly correspond. (Notice that the derivation starting with the factual premise of our discursivity does not make it therefore empirical. The derivation takes place a priori from the laws of our discursive capacity for thinking, *given* that we have such a capacity, and no other.)

This is why Kant, in §16 of TD, starts with the famous proposition “The I think *must be able* to accompany all my representations”, which is the principle of discursive thought, or indeed the principle of apperception.³ The dense, some might say convoluted, argument that ensues in the next paragraphs in §§16 and 17 contains, in my view, the argument for the logical derivation of the twelve categories, *from* apperception, which are thus shown to be the necessary and formally sufficient conditions for the possibility of discursive thought in general and hence also of the thought of an object, that is, of what enables us, as discursive minds, to think or conceive of an object at all. By showing exactly how each and every category is effectively derivable from the ‘I think’-proposition, and thus constitutes the capacity for discursive thought, Kant, so I argued, can show that these *subjective* conditions of thought have *objective* validity (A89/B122); in other words, the derivation argument shows that the same set of conditions, or functions of thought, governs the possibility of both subjective thought and thought of an object—these functions being the categories, when specifically referring to the *objects* of thought.

Put succinctly, my claim is that TD demonstrates, by way of the derivation argument—which is an ostensive proof in the sense that Kant indicates at A789/B817—that there is no discrepancy between what makes it possible to think in general and what makes it possible to think *of an object*. In fact, I believe it is only if we read TD in terms of the derivation argument that we can really understand how Kant is able to

show that thought itself is intrinsically objectively valid, that subjectivity itself is constitutive of objectivity, and that therefore the categories apply to the objects of experience. By showing that Kant's argument for what, following Henry Allison, I call the 'reciprocity thesis', stands up to scrutiny, I go against persistent strands of criticism of the validity of this thesis, which is the central thesis of the 'first step' of the B-Deduction, and I would argue of the Deduction as a whole.⁴ The criticism namely is that Kant does not account, among other things, for the difference between a claim that says that self-consciousness, and so the subjective conditions or functions of thought, are necessary for the thought of an object, and the ostensibly different claim that self-consciousness is *sufficient* for such a thought of an object. Hence, it is argued that there is an unbridgeable gap between the two claims that invalidates the main claim of there being an entailment relation between the subjective conditions of thought and the categories as the objective conditions of experience, which Kant proposes in §§16 and 17. In KDA, I explained, in great detail, that this criticism is based on a false, i.e. psychological, reading of transcendental self-consciousness, an inflated interpretation of the scope of the analyticity of apperception, and a misapprehension of the relation between the analytic unity of consciousness and the original-synthetic unity of apperception.

These are complex issues, which of course I cannot rehearse here in the detail with which I addressed them in KDA. But they are clearly important for evaluating the success of TD. I go over the main points concerning the scope of the principle of apperception in Chap. 4, when I rebut the charges of a gap in Kant's central argument about the relation between self-consciousness and objectivity. If Kant's critics, regarding the reciprocity thesis, are right, TD must be considered a failure, given how central this claim is to TD. One might perhaps want to argue that in TD Kant does not argue for the sufficient conditions of objective thought (and, if we take in the 'second step' as well, the conditions of objective *experience*), but merely for the very general thesis that the categories are the *necessary* conditions of objective experience, that is, that the categories are necessarily applicable to objective experience.⁵ But these commentators—if not the majority of readers of TD—confuse Kant's argument in the Analogies with the one in TD

(hence, they often hastily turn to the Analogies for the specifics regarding what constitutes categorially determined experience).⁶ In TD, Kant must show *how* and not just *that* the categories are necessarily applicable to objects of experience. The 'how' provides insight into the manner in which the subject of thought and the object of thought are necessarily a priori linked, and how each of the twelve categories is an a priori constitutive element in and of this connection. A detailed account of the mutual implicatedness, or what I called the 'rigorous coextensivity', of the synthetic and analytic aspects of transcendental apperception shows this. The exposition of the a priori grounds of experience in TD is philosophically more fundamental, and thus more general, than the account of the principles of experience offered in the Analogies. To suggest that we need the Analogies to understand TD thus rests on a misunderstanding of the order of fundamentality of the different sections in the *Critique*.

In KDA, I also assessed the question of the sense in which we must read Kant's distinction, in §13 of TD, between the *quid juris* and a *quid facti*, and what Kant in fact means by a transcendental deduction of the categories. I looked at arguments provided in this context by Henrich (1989), Longuenesse (1998), Proops (2003), Seeberg (2006), and most recently Callanan (2011). I particularly criticised Longuenesse and Callanan for blurring the distinction between the transcendental and empirical orders in Kant's reasoning. I also considered the vexed interpretative issue whether the structure of the argument of TD is either regressive or progressive. Most commentators hold either of the two possibilities to reflect accurately the structure of Kant's argument in the B-Deduction. I argued that the argument of TD (either in the A- or B-version) is both, and necessarily so⁷; this view ties in seamlessly with my claim concerning the a priori derivation of the categories from the 'I think'.

In the present chapter, I consider critical arguments levelled against central elements of my book, in particular, my claim that TD is a logical deduction or derivation *from* apperception. My critics raise several key points in regard to my reading of TD in KDA.⁸ But, as I shall argue, some of these criticisms also betray persistent misconceptions about TD. In responding, I shall try and address the diverse

objections in as much detail as I can within the space available. This will also give me the opportunity to spell out again the main planks of my interpretation in KDA. In response to Dyck (2014), I explain that there are good reasons to believe that the deduction of the categories must indeed be seen as a *logical* derivation from the unity of apperception, and also why this view of TD does not make MD redundant, as Dyck contends (Sect. 2.2). Furthermore, I maintain that the analytic unity of consciousness is crucially important to the argument of §16 of TD (Sect. 2.3). Thirdly, I argue that the categories are more intimately related to the functions of judgement than some interpreters, such as Dyck, make them out to be. I discuss the example that Dyck takes, i.e. the category of substance (Sect. 2.4).

In response to Stephenson (2014), I address his criticism of how I construe the reciprocity thesis, and reply to his charge that I did not properly tackle the problem of the so-called gap that some authors have argued weakens Kant's argument in the 'first step' of TD (Sects. 2.5, 2.6). I also defend my claim that the derivation of the categories is a proper deduction, by answering Stephenson's critique of a level confusion in my argument and by pointing out why he is mistaken in thinking that showing that the categories apply to objects of experience is not entailed by showing that the categories are instantiated in the experience of objects (Sect. 2.7), a topic I take up again in Chap. 4.

Quarfood (2014) raises a genuine problem for my interpretation of how contingency, the negative counterpart of the modal category of necessity, and something that represents what is utterly a posteriori, is to be derived a priori from apperception. I attempt to formulate the beginning of an answer to this dilemma regarding contingency in Sect. 2.8.

First, however, I should like to make a general comment. Both Dyck and Stephenson characterise the progressive argument of TD as having to do with refuting the sceptic, in contrast to a regressive construal of TD that is not going to cajole the sceptic into accepting defeat. And they appear to think that I read the progressive argument as such. However, I never claim that the progressive argument in TD is meant to refute the sceptic and nor do I believe that it should. On the contrary, I argue that on my reading the progressive argument concerns the need, for Kant, to provide a 'dogmatic' proof of sorts of the epistemological

claims that he is making in TD, something that a mere regressive argument cannot accomplish. As I pointed out in KDA (Chap. 4), the progressive argument is indeed often construed in terms of an anti-sceptical argument, but, as I also made it clear, I agree with Ameriks (1978, 2003, 2005) that Kant does not *intend* with TD to provide a knock-down argument against the sceptic.

While I would not deny that in some global sense Kant is responding to (Humean) scepticism, it is striking that some commentators continue to read TD as if Kant were *exclusively* or *primarily* formulating a response to the sceptic. However, the central argument in TD, on my interpretation, is *not* an argument against the sceptic but rather a demonstration of how a priori concepts, the stock-in-trade of metaphysics, can be seen to justifiably apply to experience (and to the *objects* of experience). This involves primarily a response to, and critique of, the rationalists, who failed to ask the question of whether metaphysical concepts are indeed used justifiably. Only by extension does it counter or invalidate, more by implication than intention, Hume's sceptical claims about the objective validity of our beliefs about objects (see in particular §26 of the B-Deduction), notwithstanding the paragraph on Locke and Hume that Kant added to §14 (the Transition to the Transcendental Deduction) in the B-edition (see B127–8), which might seem to bolster the idea that TD is designed as a response to the sceptic. I shall again clarify my position on the supposed anti-sceptical thrust of TD in my response to Stephenson (Sect. 2.5), who dwells on the issue of scepticism in connection with the so-called reciprocity thesis, which is central to what, borrowing a phrase from Pereboom (2001), I called the “master argument” of TD in KDA.

2.2 Why the Deduction Is a Logical Derivation from Apperception

Dyck raises the following specific objections to central elements of my interpretation: (1) He (a) questions the “textual basis” of my claim, based on B142, that Kant indeed means to *literally* ‘derive’ the categories

from the unity of apperception, and (b) asks whether my account of the derivation of the categories does not threaten to “replace” or “make redundant” “any effort to derive the categories by means of the forms of judgment” (Dyck 2014:69, 71); (2) he objects that I do not sufficiently distinguish between “the consciousness of the single, identical subject of thought” (Dyck 2014:73) and the analytical unity of consciousness or apperception, and, by implication, questions the tenability of my claim about the ‘rigorous coextensivity’ between the synthetic and analytic unities of apperception; and (3) by raising some doubts about the derivation of the category ‘substance’ in particular, he questions whether the derivation from apperception can indeed be of the *categories*, rather than merely of the *logical functions of thought*. Let me address these points in turn. I address point 1 in Sects. 2.2.1 and 2.2.2 and points 2 and 3 in Sects. 2.3 and 2.4 respectively.

2.2.1 Why the Deduction Is a Logical Derivation from Apperception: The Rationale for My Interpretation

Dyck questions my interpretation of B142 as a basis for the idea—one of the central theses of my book—that the categories are derived from the unity of apperception or thought in the sense that the categories can be deduced (strictly) logically from the unity of thought as a premise in an argument (ad 1a). He believes that I should have considered alternative readings, readings that take ‘derivation’ in a looser sense, that is, in a sense different from *logical* deduction, because Kant himself uses the term in different senses at e.g. B238/A193, B140, and B127–8, where it seems clear that ‘derivation’ cannot be taken to mean (strict) logical derivation from a premise in an argument.⁹

First, my thesis that the categories are derivable from apperception does not hinge on the correct interpretation of B142 as such, or the correct interpretation of the term ‘derive’ or its cognate ‘derivation’ (Kant’s *ableiten* and *Ableitung* respectively). B142 textually supports my reading. Evidently, the term ‘derivation’ can have variant meanings in different contexts, but mostly Kant just means ‘deduction’ in the

standard sense (see further the discussion in KDA, Chap. 3). Secondly, a parallel passage in *Prolegomena* §39 (Prol, 4:322), where Kant expounds on “the system of categories” and its deduction, pretty clearly leaves no other reasonable option than to read ‘derivation’ in purely logical terms, especially if we take the context of §39 into account—basically, this section is the *Prolegomena*’s version of MD. Notice that, a bit later in that section, Kant in fact equates *Ableitung* and *Deduktion* (Prol, 4:324.31–2; cf. KDA, pp. 226–227n.5). In the passage in §39, Kant writes:

Nothing can be more desirable to a philosopher than to be able to derive, *a priori* from one principle, the multiplicity of concepts or basic principles that previously had exhibited themselves to him piecemeal in the use he had made of them *in concreto*, and in this way to be able to unite them all in one cognition. (Prol, 4:322 [Kant 2002:114])

Here, Kant contrasts two ways of exhibiting the pure concepts: either by *a priori* derivation from a principle, or by a posteriori gradual aggregation. The latter is not a viable way to proceed for Kant. Kant further explains, in this section, that the functions of the understanding can be “fully surveyed”, and that the “pure concepts” “arise” from them, “determined *exhaustively* and *with precision*” (Prol, 4:323 [Kant 2002:115]; emphasis added). What Kant means by this is that the “kind of cognition” that we are after in an analysis of pure concepts is not a loose “aggregate” of concepts, but a “division” whose “necessity” we comprehend, as in a “system” (Prol, 4:322 [Kant 2002:114]), “founded on a universal principle”, and which “forms a closed circle” (Prol, 4:325–326 [Kant 2002:117]), in which each and every part is reciprocally integrated.¹⁰ In the introduction to the Clue sections in the *Critique*, Kant similarly writes:

Transcendental philosophy has the advantage but also the obligation to seek its concepts in accordance with a principle, since they spring pure and unmixed from the understanding, as absolute unity, and must therefore be connected among themselves in accordance with a concept or idea. Such a connection, however, provides a rule by means of which the

place of *each* pure concept of the understanding and the *completeness* of all of them together can be *determined a priori*, which would otherwise depend upon whim or chance. (A67/B92; emphasis added)

This rule for the a priori determination of each of the pure concepts lies in the act of the understanding as a capacity to judge (see Prol, 4:323). The derivation consists in a step-by-step demonstration—by virtue of an “*analysis of the faculty of understanding*” (A65/B90)—of the way in which each single pure concept, or category, forms an integral part of the nature of the discursive understanding, and how they all hang together systematically as jointly constituting the capacity to think, which for Kant is a capacity to judge. Each of the twelve categories corresponds to or is identical to each of the twelve functions of the understanding (or judgement), and so each category “analytically pertains”, as I put it in KDA, to the unity of thought (on the identity relation between function of thought and category, see further below in response to Dyck’s third objection).

It is in this way that I claim, as I did in KDA, that the categories can all be derived *from* the ‘I think’, *from* apperception as being the capacity to think, since together they constitute the unity of thought, which Kant claims is intrinsically *objectively* valid. A different way to describe the close relation between the categories and the unity of thought or the unity of apperception would be to emphasise the “conformity” between the categories and the subjective conditions of thought, which are their grounds, precisely as Kant asserts this in his later essay *On a Discovery*, where he says that the original acquisition of the categories “presupposes nothing innate save the subjective conditions of the spontaneity of thought (conformity with the unity of apperception [*Gemäßheit mit der Einheit der Apperception*])” (ÜE, 8:223 [Kant 2002:313; trans. emended]). I take this to confirm my view that the categories are a priori derivable from, or “analytically pertain to”, apperception.

Ironically, this way of reading the derivation of the categories is not necessarily at odds with what Dyck suggests, at the end of his account, is an *alternative* reading of the derivation question, namely that “the categories might be understood to be ‘derived’ from apperception only

insofar as they find the original source of their unity in the identity of the thinking subject" (2014:75). It is not clear though how Dyck himself envisages how such a derivation would work and how categories "find" their source of unity in the identity of the subject, if not in the sense in which I take it to be, namely in terms of a logical, a priori derivation from apperception. Notice the connotation of arbitrariness conveyed by the term 'find' that Dyck employs, as if the categories come together haphazardly, quite clearly the contrary of what Kant is aiming for.¹¹ As I explained in detail in Chap. 3 of KDA, where I delineate in which sense Kant's deduction is not *merely* a justification modelled after a *Deduktionsschrift*, Kant's philosophical notion of deduction is primarily one for which ostensive proof is the model. Ostensive proofs are genetic proofs, which reveal "not only truth but also at the same time its *genesis*, its generative source" (V-Lo/Blomberg, 24:233 [Kant 1992:186]), and transcendental proofs are such ostensive proofs (A794/B822) (see further KDA, pp. 26–27). I cannot see how on Dyck's alternative, looser reading of derivation, Kant's deduction could be seen indeed to *show*, ostensively, how the categories "spring pure and unmixed from the understanding" (A67/B92).¹²

Thus, thirdly, it is a fortiori unclear how Dyck thinks that by 'derivation' Kant does *not* mean a logical derivation, but rather, by using the term "assert[s] that the categories are *merely dependent on* the 'I think' as, for instance, the original source of their unity" (2014: 70–71; emphasis added). What can be meant by mere dependence? In what sense do the categories depend on the 'I think'? Are they logically dependent on it? Well, apparently not, given the "merely" that Dyck adds. Or are they psychologically, empirically dependent? It is not clear. Moreover, how does Dyck think the relation between the categories and the original source of their unity (which, moreover, is *what* on his view?) can be demonstrated or shown if not by way of an ostensive proof? At least *some* sort of reasoning from premises is required to show the dependence. The categories cannot be plucked out of thin air!¹³ Dyck's proposal is not really a viable alternative reading of B142, but seems rather to stem from the fact that he is at a loss to explain Kant's claim here.

Such a lack of clarity in regard to the derivation question is quite standard in the literature. No single scholar of Kant has been able to explain how precisely the account of self-consciousness hangs together with the categories as the transcendental conditions of experience, precisely because nobody—excepting perhaps Henrich (1976),¹⁴ and to an extent Fichte and Kant’s early commentator J.H. Tieftrunk, as I explain in KDA—has been able to explain the putative dependency relation between self-consciousness and the categories. Only by explaining this relation—and not just stipulating that there is one—between, on the one hand, self-consciousness, or apperception, which is the principle of discursive thought, and, on the other, the categories as the enabling conditions of the object of experience as well as the experience of the object, can one explain *how* and *why*, and not just *that*, the categories apply to experience.

2.2.2 Does My Reading of TD “Threaten to Replace” MD?

There is another apparent issue concerning the derivation question that bothers Dyck (ad 1b). He thinks that, contrary to what I declare in KDA, my reading of the derivation of the categories from apperception “threatens to replace” or “make redundant” “any effort to derive the categories by means of the forms of judgment” (Dyck 2014:71), that is, on my reading of TD there would no longer be a need for MD. I am unclear about what he means by a derivation of the categories “by means of” the forms of judgement. One would think that in MD, in §10 (the Third Clue section), the categories are derived *from* the forms of judgement, or more precisely, from the *functions* of thought in judgement. But at any rate, by no means is the metaphysical deduction of the categories from the functions of judgement made redundant on my reading of the derivation of the categories *from apperception*, which is the project of TD (or at least its first part in the B-version). And I made this quite clear in KDA. I argued that MD is merely the first step in the overall proof of the deduction of the categories from the functions of

thought. (Some have argued that MD addresses the *quid facti*, whereas TD is concerned with the *quid juris*, but I do not think this is right.¹⁵)

What Kant does in MD, especially in the first two sections (the introductory section and §9), which are the two sections that Reich (2001) and Wolff (1995) concentrate on, is just to classify, in a preliminary fashion, the elementary functions of thought in a table of judgement based on the idea that the capacity to think, and thus to know, is identical to the capacity to judge, i.e. the understanding. There is a reason why Kant talks, in these sections in MD, about *clues* to finding the categories. There is no (ostensive) proof yet of the intimate link between the functions of thought and the categories, also not in §10, where he provides the table of the categories, based on the table of judgement. It is first in TD that Kant shows, non-question-beggingly, how the capacity for thinking is a capacity for objectively valid judging by means of the categories, by deriving the categories and the general form of judgement *from* apperception, thus proving the correspondence between the table of judgement and the table of the categories. In contrast to MD, where Kant just *says* that the understanding is a capacity for judging, in TD he *shows* it (whether he does so successfully is of course another question).

Hence, as I made it quite clear in KDA, my reading of TD does *not* conflict with Reich's, Wolff's, and also Longuenesse's (2006) reading of MD, who in contrast to my focus concentrate on the logical derivation of the *functions of judgement* in the First Clue section (A67–9/B92–4), culminating in the table of *judgement* in the Second Clue section (§9). When I talk about a “reappraisal” of the famous *Leitfaden* passage in the Third Clue section, at B104–5 (KDA, p. 3), I do not mean to say—Dyck misreads this—that *Kant's account of TD* is a “reappraisal” of MD, as if it needed revising, but rather that on my reading of TD a “reappraisal” of the *Leitfaden* is warranted in light of the fact that, in my view, the standard reading does not fully appreciate what is at stake (and I provide such a reappraisal in Chap. 5 of KDA). The derivation of the categories of TD is thus certainly not “an alternative to Kant's analysis in §§9–10” (Dyck 2014:71, corrected¹⁶), and I did not portray it as such.

2.3 The Role of the Analytic Unity of Consciousness

Let me now get to Dyck's second central objection (ad 2). He believes that I "overlook important differences between the analytic unity of apperception as evinced in common concepts and the consciousness of the unity or identity of the subject" (Dyck 2014:72). According to Dyck, unlike the consciousness that accompanies the analytic unity of a common concept, of the concept <red>, say, that is attributed to a manifold of representations of red objects, the consciousness of "the identity of the subject as such is clearly not an element *in* the content of the manifold of representations (i.e. a partial representation); it is, as Kant stresses, that *to which* the entire manifold of representations belongs (B131–132)" (2014:72). Therefore, Dyck believes that "the consciousness of the single, identical subject of thought is not aptly characterized in terms of an analytical unity of consciousness" (2014:73). This strikes me as an odd take on Kant's argument in §16, for several reasons, not least because it threatens to undermine the central claim of the Deduction, namely that *the same function* of the understanding unites both concepts, *by means of an analytic unity*, and intuition, *by means of a synthetic unity*, in a judgement (B104–5), which I take to refer to the analytic and synthetic unities of apperception respectively underpinning the one function of the understanding. It also dissents from the majority view on this central issue.

The passage in B134 directly following the passage at issue here (namely, the passage to which the footnote concerning the consciousness of an analytic unity in conceptual representations is appended, i.e. B133¹⁷) makes it clear that the belonging of representations "to me", i.e. to me as identical subject, comes down to nothing but an act of the unification of one's representations "in a self-consciousness", that is, only by "comprehend[ing] their manifold in a consciousness do I call them all together *my* representations" (B134). Representations belong to me in that I take them together as belonging to me. And it is the synthetic unity that results from this act of comprehending that first establishes, or, is "the ground of the identity of apperception itself" (B134).

The identical subject across the representations that are accompanied by the 'I think' is thus established *in* the act of synthesis. I do not see how Dyck can argue that a strict distinction should be heeded between, on the one hand, the analytic unity of consciousness—that is, the identical representation contained *in* the manifold of representations that I regard as *mine*, all of the different representations <tomato>, <bus 93 to Putney> and <Labour>, say, which I represent when I think of the concept <red>, the identical representation here being the indexical 'I' that, as common mark, is contained in all of *my* representations, qua their being my representations, as much as the partial representation <red> that is contained in all of those representations <tomato> etc.—and, on the other hand, the identity “to which”, as Dyck emphasises, these representations belong.¹⁸ Are the 'I', *to which* my representations belong, and my representations themselves, as collective manifold *in* which an identical indexical 'I' is contained that they all share in order to be *my* representations, to be seen as distinct, as Dyck seems to be saying? In which sense? And on the basis of what? And if so, what connects *my* collective representations to the 'I' if it is not through the a priori synthesis by means of which I synthesise my representations *as* first belonging to me?

Of course, I am not logically compelled to represent representations that are *conceptually* analytically related: I can think a manifold of representations that other than being concepts have nothing (prima facie at least) in common, <hoarfrost>, <Heidegger> and <haggis>, say. However, whatever the case may be, as represented concepts they would still share the one partial representation that they must have in common in order to be represented *by me* at all, namely the partial representation 'I', which is the same in all representations that I (conjointly or not) represent. As Kant puts it, “concepts [are] possible through apperception [...] [i]n that *I* represent to myself the identity of *my* apperception *in* many representations” (V-Met/Mron, 29:889 [Kant 2001:257]; emphasis added).¹⁹ This holds in the first instance for how single concepts are formed, but also for the representation of series of concepts or the representation of *relations* between concepts, such as the relation between <hoarfrost>, <Heidegger>, and <haggis> in e.g. my judgement

“Heidegger ate haggis while gazing in astonishment at the hoarfrost around his cabin” (suppose that I was invited to Todtnauberg).

The analytic unity of consciousness, which for its ground rests on the act of (synthetically) comprehending a manifold of representations, thus constitutes the identity of my subject as the subject of that complex thought, which consists of the multifarious representations <hoarfrost>, <Heidegger> and <haggis> (among others) in the judgement “Heidegger ate haggis...”. The ‘I’ here is the partial representation that is formally contained in all of those representations that I conjointly represent, to the extent *and to the extent only* that they are *my* representations, namely to the extent that I am the judger of that judgement. No identity of subject exists without an analytic unity of consciousness comprising all of the representations that *I* represent to the extent that those representations share the common mark ‘I think’ by being accompanied by it, and there is no analytic unity of consciousness which does not entail an identical subject of thought. As Allison (1996:58) correctly says with reference to A108, the “mind can think its identity only insofar as it can become conscious of the identity of its function or act”, namely, the act of synthesis, which underlies any analytic unity of consciousness (B133). The question could then arise as to whether there is something more about the *identity* of the subject over above the analytic *unity* of consciousness, which is constituted by the identity of the mind’s function of synthesis. I suspect that Dyck believes that there is, but it goes against Kantian doctrine. I concur fully with Allison’s position on this point:

Kant’s view is not that we require a distinct capacity to become conscious of our identity as thinking subject in order to be able to think the necessary unity of appearances; it is rather that to think such a unity is, *at the same time*, to think one’s identity. [...] I must in the very same act [of representing to myself an objective connection] take myself as an identical thinking subject. (Allison 1996:60)

Dyck’s (2014:73) claim that the analytic unity of apperception is inessential (!) to the argument of §16 of TD also throws a monkey-wrench into the debate on the self-ascription of representations that has

evolved from a crucial reading of Kant's principle of apperception, namely the austere interpretation of apperception espoused by P.F. Strawson (1968) and his acolytes, which wants to do away with synthesis altogether. Abstracting here from his (Strawson's) unfounded criticism against a priori synthesis as a piece of unwarranted psychology, I think that one of the strong aspects of Strawson's interpretation, and which certainly has been very influential in Anglophone Kant scholarship, precisely concerns how he regards the principle of apperception as an analytic principle of self-ascription of one's own representations that states that a reflexive, *analytical* relation ("a direct analytical connexion", as Strawson [1968:96] puts it) obtains between an identical 'I' and all of its (my) own representations that are self-ascribed by that 'I'. For all that Strawson is wrong about a priori synthesis, he is certainly right (*pace* Dyck) that the principle of apperception that is central to the argument of §16 is an *analytic* principle that expresses an *analytic* unity between representations insofar as they are self-ascribed by the self-same, identical subject. (In Chap. 4, in particular Sects. 4.8 and 4.9, I elaborate on the analysis of the analytic unity of apperception.)

2.4 The Intimacy Between the Categories and the Functions of Thought

Lastly, Dyck criticises my claim that categories are the functions of thought and that they therefore can be derived from the 'I' (ad 3). He takes 'substance' as an example. That "the category of substance should have its source in pure apperception", Dyck (2014:74) acknowledges, is indeed claimed by Kant in his pre-Critical work, but, as Dyck counters, this view does not survive the Critical turn. Incidentally, Dyck falsely states that I do not cite pre-Critical texts in my account of the relation between substance and apperception. It appears that he overlooked the passage in Sect. 7.1 of KDA, where I explicitly quote and briefly discuss *Reflexionen* 4674 and 4676 from the *Duisburg Nachlass* (KDA, pp. 127–128). Furthermore, in a note to the passage in which I explicitly claim that the concept of substance is conveyed by the very notion of the 'I think' as the original representation that accompanies all my

representations but cannot itself be represented by an even more original representation (KDA, p. 135), I, too, quote the very *Reflexion* that Dyck suggests is representative of the pre-Critical view I defend, namely *Reflexion* 3921 (see KDA, pp. 261–262n.12). In this *Reflexion*, Kant says that the “idea of substance actually comes from the *repraesentatione sui ipsius* [representation of oneself]” (Refl, 17:346 [Kant 2005:95]), and it indeed supports, as Dyck notes, my reading of the intimate connection between the category of substance and apperception.

More important, though not nearly as significant as he thinks, is Dyck’s criticism that in the *Critique*, as a result of the paralogisms, Kant is no longer licensed to make the claim that substance is directly derivable from the ‘I’. Only the logical function, “not yet (or no longer) the full-fledged category of substance”, says Dyck, “might be derived from the ‘I think’” (2014:74). Quoting from the A-Paralogisms, Dyck points out that “the concept of substance is used only as a function of synthesis” (A356) if it is not related to an empirical intuition; I would add to this that the mere concept of substance, *sans* empirical intuition, is used only as a function of *intellectual* synthesis (cf. B150; A401; KDA, pp. 93, 95). Dyck argues that since, for the Critical Kant, the self is no longer immediately given (in intuition) as an object, something the pre-Critical Kant putatively still believed, the basis for claiming, as I do, that the concept of substance can be derived from the ‘I’ is removed. I do not see why this would follow. The fact that, for the Critical Kant, the self can no longer be seen to be *given* as an object, or substance, in empirical intuition, leaves untouched the view that the *concept* or category of substance can be derived from the ‘I’, precisely because it represents the mere logical function of the subject that must be represented as original and “unchanging” (A107), as one of the functions of intellectual synthesis. And this is in fact what Kant says himself: it is perfectly legitimate to call the ‘I’ a substance, and hence to grant the “pure category” subjective significance, while at the same time denying it “objective significance” (A348–9; cf. A350–1; B343/A287; see further KDA, pp. 131–132).²⁰ I am not clear about what Dyck means by “full-fledged category” (2014:74), but I suppose he means ‘category under which an empirical intuition is subsumed’, a category which refers to an actually existing spatiotemporal object, in other words, a schematised

or 'objectively significant' category; if so, his criticism begs the question against my claim that the category is derivable from the 'I', since I clearly do not claim that the 'category under which an empirical intuition is subsumed', let alone a categorially determined spatiotemporal object, is so derivable.

As I argued in my book (KDA, p. 135), one way to look at the difference between the pre-Critical and Critical accounts of the relation between substance (and *mutatis mutandis* any one category) and the subject is to note that in the Paralogisms Kant specifically argues, as part of his general argument against the rationalist, that any possible claim to the substantiality of the 'I' *always already* presupposes the 'I' as the vehicle of my thoughts, and *eo ipso* also presupposes the applicability of the *concept* of substance, as one of the elementary characteristics of the subject of thought, which the rationalist then (illicitly) translates into the 'I' as being a really subsisting thing. In other words, regardless of the view one holds on the soul's substantiality, the *concept* of substance is prior to any application of it to a putatively enduring object, be it an empirical spatiotemporal object, for which an additional sensible intuition is needed, or indeed a noumenal subsisting self! Hence, the category as *pure* concept pertains to thought itself, as one of the elementary concepts that constitute it²¹; for the *concept* of substance is nothing but the function of how a subsisting subject relates to its changing predicates as its accidents, regardless of the question whether this subject *is* in fact a noumenal substance (cf. V-Met/Volck, 28:429; V-Met-L₁/Pölit, 28:239–240).

It is important to be mindful of the fact that, in Kant's view, the rationalist is not at all mistaken to associate the *concept* of substance with the (ultimate) subject of thought,²² but rather to conclude from their logical correspondence (and the given empirical experience of the self)²³ that we have indeed knowledge of the subject as a noumenal substance, as a noumenally enduring object (cf. B407). I think this insight is crucial for a proper, nuanced understanding of the relation between the pre-Critical and Critical Kant, not least with regard to the paralogisms that played a crucial role in the transition to the Critical account.²⁴

Furthermore, Kant does explicitly assert that, notwithstanding their *formal* differentiation, the categories are nothing but logical functions, i.e. “moments of the understanding” in judgements, and “will come out exactly parallel to them” (Prol, 4:302 [Kant 2002:96]; trans. emended; cf. Prol 4:305, 324; ÜE, 8:223; Refl 5854, 18:369–370). Certainly, the categories and the logical functions of thought are formally differentiated for the purposes of expounding the transcendental logic, but they are not *distinct* functions of the understanding. If they were distinct functions, this would run counter to the central claim of the *Leitfaden*, namely that the categories *are* the functions of thought under a different aspect, namely insofar as thought is *about* objects. Therefore, Kant concludes the ‘first step’ of TD by saying that “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” (B143).

Not coincidentally, Dyck’s denial that the categories are in fact the logical functions of thought, *sans* the intuition that synthetically connects them to real objects (cf. ÜE, 8:223), coincides with his aforementioned denial of what in KDA I labelled the “rigorous coextensivity” between the analytic and synthetic unities of apperception. However, without that coextensivity, not much of Kant’s main claim in the *Leitfaden* about the parallelism of the logical functions of thought which combine concepts by means of analytic unity and the categories as the concepts providing the synthetic a priori content of judgement—which is in fact the main claim behind the entire project of TD—remains intelligible, inasmuch as the logical functions of thought are purely *analytical* rules for thought, which like intuitions of objects rest on an original synthetic unity of apperception. To say that those logical functions of thought, as analytical rules for combining representations into concepts “by means of analytic unity”, are not central to the argument of TD, as Dyck appears to claim, seems to me to miss the pivotal point of Kant’s deduction project.

All things considered, Dyck’s concluding assertion that “any putative derivation of the *category* of substance (as opposed to the mere logical function) from pure apperception is doomed to be unsuccessful” (2014:74) is rather exaggerated, to say the least.

2.5 The Reciprocity Thesis and the Progressive Argument

Fortunately, Stephenson (2014) agrees with my take on TD as (a) concerning the 'en bloc' applicability of the categories to experience, which should not be confused with the mistaken, but popular, view that in TD Kant argues 'en bloc', as it were, that the categories necessarily apply to experience, rather than argues for *each* of the categories as so applicable; and as (b) pivoting on what Allison has called 'the reciprocity thesis', namely the claim that the "unity of consciousness is necessary and sufficient for objectivity" (KDA, p. 53), which is often disputed, since it is not *prima facie* clear why one would agree with the claim that in addition to being the necessary condition of the experience of objects, the unity of consciousness is also sufficient for the experience of objects, let alone for the existence of the objects of experience.

However, Stephenson criticises the way I have argued for these claims. His incisive critique gives me the opportunity to clarify my position and meanwhile point out why I think Stephenson's own construal of TD is mistaken. It turns out that he endorses the reciprocity thesis in a way fundamentally different from how I read it, so it remains to be seen how much agreement between our interpretations there in fact is. This difference hinges on how we interpret the meaning of Kant's term 'knowledge' (*Erkenntnis*) and, relatedly, 'object' and 'objectivity', differently (see further below); but also, how we regard the relation between the subject or self-consciousness and object or object-consciousness, which Stephenson appears to see in a way that I would say looks like too short an argument to reciprocity. The reason why I suspect he sees problems with my reading is because he appears to take a traditionally 'realist' conception of object and of knowledge for granted when interpreting TD, which of course should *not* be taken for granted, given the nature of Kant's Critical turn in philosophy. Let me address Stephenson's criticisms in the order he raises them: (1) the reciprocity thesis and Stephenson's criticisms of my construal of it, and (2) his claim that my proposed derivation of the categories is not actually a deduction at all, and fails to show how the categories are applied to objects.

Ad 1: Stephenson portrays a progressive argument, in the standard sense, in terms of an argument that works against the sceptic, namely an “argument a [...] in the face of sceptic s ”, which meets the two conditions, as Stephenson says, that “the basic premise of a is accepted by s ” and “the conclusion of a is rejected by s ” (2014:78). That is undoubtedly true on one reading of what a progressive argument amounts to, such as is the case for example in Strawson’s take on TD. And I agree with Stephenson’s general account of how such anti-sceptical or transcendental arguments work. But as I argued in KDA (Chap. 4), that is certainly not the only way to construe a progressive argument and nor, I believe, the way that Kant wants his argument in TD to be understood. Stephenson suggests that on the progressive construal a “transcendental argument” can be seen as “*refut[ing]* the sceptic who denies the existence of the external world” (2014:78) through showing that the latter is in fact the necessary condition of having a unitary consciousness at all. But this looks more like the concern of the Refutation of Idealism,²⁵ not of TD, which operates on a much abstracter level, where the argument is about what it means to have a conception of objectivity in the most general terms. External world problems are the least of Kant’s worries in TD.

Nevertheless, to argue that it is a general conception of objectivity that is at stake, as I do in my book, does not imply, as Stephenson suggests, that on such a reading Kant would only be committed to a theory that argues that “experience at least *seems* to present us with an independent world of causally interacting objects”, but that experience does not “necessarily [represent objects] veridically” (2014:79). The general conception of objectivity at stake in TD is not merely about how *we conceive of* or *experience* objects, but also about how, given the constraints of discursive cognition, objects can be veridically represented at all, and so how empirical objects are constituted qua their being objects (cf. B138), given that there are such objects, and that we make knowledge claims about them.

So in what sense do I think that TD is still a progressive argument if it is not construed in the standard sceptic-refuting sense? Stephenson’s reconstruction of my take on the dual progressive-regressive nature of TD, though splendidly succinct, is formally to the point.

The progressive argument on my reading starts from the unity of consciousness and moves to objective experience, or to be more precise, to *possible* experience, since the argument is entirely a priori and does not concern questions of how we actually have instantiations of experience (*quaestiones facti*, which lie outside the remit of TD). Stephenson then asks: “But who is the sceptical target here?” (2014:78). Answer: Nobody is! Or at least, there is no *explicit* sceptical target. The progressive argument in TD as I construe it is just to show, in a priori manner, *how* it is possible that we have knowledge of the object given that we do make claims about it and that there *are* such objects as we make claims about. This kind of progressive argument is not in the least meant to refute the sceptic. The progressive nature of the argument has merely to do with the a priori nature of the proof that delineates the how-possible question.

2.6 Putative Gaps in Kant's Argument: A Preliminary Account

Stephenson's reading of the progressive argument as anti-sceptical ties in with how he reads the reciprocity thesis and the supposed gap in Kant's reasoning. However, in KDA, I'm talking about a different gap from Stephenson's. It is not the gap between, on the one hand, how we think about objects and, on the other hand, *really existing* objects out there. Kant is not worried about the fact that our experience is about such objects, nor is he worried about whether there are such objects in the first place. His question concerns how it is *possible* that we can have knowledge of them, which means how we can have a priori knowledge of them, if the knowledge is to be more than accumulated empirical knowledge—put differently, the main question concerns the way or ways in which we understand something to be an object for us, what constitutes an object's *objectivity*. It may be of course that some readers of TD think that Kant needs first to prove that there are objects or that our experience is about those objects; but I'd say those readers are fundamentally mistaken about Kant's epistemic confidence. The gap that

I discuss in KDA (in Chap. 4) is the perceived gap that Guyer (1992), Mohr (1991), Pereboom (2001), Carl (1989a, b) and others²⁶ talk about, namely the gap between the principle of apperception, which is analytic, and the *objective* unity of apperception, which clearly is not analytic. Let's call this Gap₁. Stephenson does not appear to see a problem there. And he is right, there isn't, as I indeed argue against Guyer et al.

By contrast, Stephenson's gap concerns the (perceived) gap between what *we experience* of objects, or that we make claims about them, and *the objects* we veridically experience, the things that we make claims about (presumably Stroud's problem).²⁷ Let's call this Gap₂. Of course, Gap₁ and Gap₂ are in some way related. They both concern the way the reciprocity thesis should be understood: does the unity of consciousness constitute a necessary *and* a sufficient condition of experience of objects? But Stephenson's gap, Gap₂, concerns an additional claim that the reciprocity thesis entails, which Stephenson does not spell out as such, namely, that the unity of consciousness is not just sufficient for the *experience* of objects but also sufficient for the *objects* of experience, that is, sufficient for the *existence* of objects (in some sense) (see KDA, pp. 50–51). Stephenson himself talks in terms of the following fourfold gap: “(i) truths about the structure of our conceptual scheme are not necessarily truths about the world, (ii) constraints on what set of beliefs is rationally coherent do not necessarily tell us anything about what set of beliefs is true; (iii) that we must apply the categories to objects does not obviously entail that the categories must apply to objects; (iv) that we must apply the categories does not mean that we are justified in doing so” (2014:79). And Stephenson claims that I have not tried to bridge this fourfold gap.

It seems to me that, in his characterisation of the gap that Kant presumably must bridge, Stephenson here misses Kant's central point, in the ‘first step’ of the B-Deduction, about the objective unity of apperception as *defining* the objective validity of a judgement, which is not just the truth value of a judgement—the truth value is merely a surface aspect of a proposition that expresses an assertion, which can be either false or true. Rather, objective validity concerns the way a judgement,

as the manner in which cognitions are brought to the objective unity of apperception, is always already intimately connected with, and so truly corresponds to, the object it is about (see further Chap. 3). There is no fundamental gap between a claim *about* the object and the object itself to be bridged. So it seems to me that by defining object as “that in the concept of which the manifold of a given intuition is *united*” (B137), and given that the “synthetic unity of consciousness is [...] an **objective** condition of all cognition, not merely something I myself need in order to cognize an object but rather something under which every intuition must stand *in order to become an object for me*” (B138; boldface mine), Kant asserts that the unity of consciousness is not just a necessary and (formally) sufficient condition of *our experience* of an object, but also the necessary and (formally) sufficient condition of the object itself, that is, of its existence (qua object of experience). In §19, Kant then asserts that a judgement’s objective validity, the *condition* of its truth, lies in the objective unity of apperception.²⁸ Kant has thus made clear that our conceptual scheme corresponds to the world, and that there is no worry about whether the truth of our conceptual scheme is true about the world or not: there is no ‘outside’ of our conceptual scheme, for the object that we represent as over against our representations is *nothing but* the unity of apperception (cf. A104–5, 107). This deals with the first two prongs of Stephenson’s description of the gap. The third prong mistakenly supposes that the categories are merely necessary conditions of the experience of objects and not the necessary conditions of the objects of experience, but Kant of course claims that the categories are both (A111; B197/A158).

All of these claims of Kant I have dealt with in my book, so it seems to me that the above three prongs of Stephenson’s account of the gap have been sufficiently answered. As to the last prong: it is clear that the application of the categories is justified once we have shown that they are necessary for the experience of objects and for the objects’ existence, as objects for us (cf. A95–8; A128–30). Any further question about justification is irrelevant. With point (iv) Stephenson seems to postulate that the necessity for the application of the categories in separation from the justifiability of doing so concerns some sort of psychological necessity, but this is of course not at issue in TD.

Now Stephenson might insist: the problem here is that you have only claimed a reciprocity between apperception and an object *in general*. Gap₂ still stands, for the gap between our experience and *real*, spatiotemporal, empirical objects has not been bridged, so that it has still not been shown that our experience is veridically *of those spatiotemporal objects*. I have admittedly not dealt specifically with that presumed gap in KDA, since the relation between apperception and spatiotemporally located objects is the topic of the 'second step' of the B-Deduction, which did not fall within the remit of my book. Stephenson might have been misled by my differentiation of the terms 'cognition' and 'knowledge' (see e.g. KDA, p. 60) into believing that I argue that Kant's argument in TD as a whole is merely about the conditions of our beliefs or claims about objects, and not about (true) knowledge. But when I made that differentiation, which corresponds to Kant's distinction between "thinking an object" and "knowing" (*erkennen*) one (B146), I was talking merely about the 'first step' of the B-Deduction in contrast to the 'second step', where Kant indeed expounds on the conditions of *knowing* a spatiotemporal object, rather than just about the conditions of *cognising* an object in general. Importantly, the unity of apperception, in its guise as the productive imagination, as an a priori effect of the understanding on sensibility itself (B151–2), is also the necessary and formally sufficient condition of *empirical knowledge of spatiotemporal objects* (I discuss this further in Chap. 7).

Notice that I say *formally* sufficient, for obviously Kant cannot claim that apperception, or the set of categories even in their schematised form, is the *materially* sufficient condition of empirical knowledge, since that would mean a conflation of the a priori and a posteriori conditions of knowledge. But there is no sense (for Kant, at least) in which there is a gap between what can be a priori determined and what is a posteriori given, which would have to be bridged to explain the possibility of empirical knowledge. Again, *transcendental* philosophy is not concerned with proving *that* there are objects that empirically affect us or *that* we have sensations that can be taken to refer to objects, nor would it be able to prove that. The a posteriori is just assumed as given: we just happen to *have* sensations, and we can be confident that there *are* objects out there of which these sensations are the causal effects. Hence, the

premise of Kant's argument on a regressive construal of TD is the givenness of the empirical objects that we experience.

There is debate as to whether Kant's *Erkenntnis* should be translated as 'cognition' (as do Guyer/Wood) rather than as 'knowledge' (as does Kemp Smith). The fact that Kant allows false *Erkenntnis* (A58/B83) would seem to indicate that it cannot be translated as 'knowledge' (in our post-Gettier contemporary sense). Moreover, Kant himself appears to identify *Erkenntnis* with the Latin *cognitio* in the *Stufenleiter* (A320/B376–7). However, we should be careful not to gloss Kant's term *Erkenntnis* as if it signified merely being about the subjective conditions of cognition or a mental act or an epistemic attitude, and not also about the *objective* conditions under we can in fact *know things* or *facts* (in the contemporary sense). While strictly speaking knowledge (*Erkenntnis*) in Kant's sense is, taken as such, not the same as knowledge in our contemporary sense, neither is it *just* cognition as a mere subjective *capacity for* knowing, which would rather be "to think of an object" on Kant's account, in contrast to knowing one (B146).²⁹ I employ, as I did in KDA, the terms 'cognition' and 'knowledge' as reflecting Kant's distinction between 'thinking an object' and 'knowing an object' respectively (B146; B165). Notice that Kant's *Erkenntnis* must be divided into its transcendental and empirical meanings, where the empirical meaning of *Erkenntnis* (i.e. *empirische Erkenntnis*) could be identified as knowledge in our contemporary sense (cf. B165–6). The transcendental meaning of *Erkenntnis*, as it is discussed in TD, concerns the transcendental or a priori conditions of such knowledge, but this should not be conflated with the psychological condition or capacity for knowing, nor with the *merely* conceptual capacity or the capacity to think or judge. The transcendental or a priori conditions of empirical knowledge, which includes pure intuition, is what *makes* empirical knowledge knowledge. It is these knowledge-making conditions that are at issue in TD. There is nothing beyond the knowledge-making conditions that would constitute knowledge, *in addition to* those conditions and given sensible input (empirical intuitions).

Concluding my reply to Stephenson's first objection, there is no ground for arguing that, given the additional argument of the 'second step', Gap₂ stands and needs resolving or that I was not aware of, or

tried to “move around” it or “put [it] to one side” (Stephenson 2014:80) in KDA. (In Chap. 4, this volume, I examine in detail further similar arguments about the putative Gap in TD, put forward by James Van Cleve and Anil Gomes.) Nor do I believe that Stephenson’s own reconstruction of the reciprocity thesis is sufficient to close any putative Gap₂. On his construal, “possible object-consciousness is a necessary condition for possible self-consciousness”, since the latter is always a “consciousness of one’s consciousness of objects”. That is, “self-consciousness, as second-order consciousness, which can be represented as C(C(O)), thus requires not only the concept of a subject but also the concept of an object, since this is also one of the things inside the brackets” (Stephenson 2014:82). It looks like Stephenson helps himself to the reciprocity between object and subject by means of a short (i.e. merely conceptual) argument that leaves out the specific categories altogether. I do not think this is an adequate account of the progressive argument of TD, and it certainly would not coax a sceptic into accepting defeat.

Moreover, this construal will not help Stephenson establish the connection between arguing that we must apply the categories and arguing that the categories apply to the objects of our experience, and not just to a *concept* of the object, the issue which he in fact accuses me of fudging (cf. his point iii in the aforementioned fourfold gap). This brings me to Stephenson’s second criticism, that is, that my attempted derivation of the categories from apperception is in fact not a deduction at all, for apparently it fails to answer the *quid juris* question. From the scant remark about this last point towards the end of his essay it is difficult to gather on what grounds Stephenson thinks that I fail to address the *quid juris* (despite devoting many pages and even a whole chapter to it in KDA), so I shall leave this particular criticism aside.

2.7 The Nature of Deduction

So to Stephenson’s criticism that my derivation is not a deduction at all (ad 2). He takes the derivation of ‘necessity’ as an example. But, unfortunately, instead of commenting on the ten pages of extensive analysis

of the categories of modality in the context of TD that precede it, he goes straight to the schematic summaries I provide in the form of argumentative steps in what I called the D-argument. Let me first note that these summary arguments are shorthand in order for me later, in the concluding Chap. 10 of KDA, to be able to expeditiously refer back to all of the D-steps; they do not replace the more extensive exposition, so they hardly present my real argument for the derivation. Secondly, Stephenson's objection reveals an intrinsic problem with any attempt to formalise Kant's transcendental proofs: on pain of loss of precision, one might end up with a formal proof that is even more unwieldy than the exposition it is supposed to translate. Perhaps, then, I should have steered clear of any quasi-formal schematisation (although I did note in a footnote that the schematisation was not meant to constitute a "syllogistically modelled inferential link of premises", but rather is just a "summary of the main arguments"; see KDA, p. 276n.46). Whatever the case may be, I find Stephenson somewhat uncharitable in attempting a bit of old school conceptual analysis in his critique of how I argue 'necessity' is to be derived from apperception.

He asks first: "[W]hat does 'pertains to' mean?" (Stephenson 2014:83). Well, exactly what *The Concise Oxford Dictionary* says it means (as one of three meanings provided): "belong to as a part or appendage or accessory", where "belong to as a part" is the most appropriate description in this context. As I suggested above (Sect. 2.2), the phrase 'pertaining to' could also be read in terms of a 'conformity' (*Gemäßheit*) between the categories and the unity of apperception, in line with Kant's assertion in response to Eberhard in ÜE, 8:223. 'Necessity' is a part of, or shows a conformity to, self-consciousness or transcendental apperception, and can thus be derived analytically from it. And so I argued with respect to all the categories.

Stephenson believes I commit an error in confusing two levels, namely in arguing "from the fact that the concept of necessity is mobilized in the statement of the theory—the analytic principle of apperception—that it therefore 'pertains to' the object of the theory—apperception itself". First, Stephenson assumes that there is a distinction, in the argument of TD, between the theory and the object of theory. But in TD the object of theory happens to be (also) the thinking

subject, who is also the subject of theorising, so that apperception itself occurs in accordance with the *principle* of apperception; the principle of apperception is not just a theory *about* apperception, rather, it is that by virtue of which apperception works. When interpretatively reconstructing the argument, presented in TD, about this thinking subject and its constitutive characteristics (i.e. the logical functions of thought with which the categories correspond), we should be careful not to artificially introduce a distinction between two levels of discourse (theory and object) on pain of losing sight of the self-reflexive aspect of Kant's reasoning. The reconstruction must retrace the cognitive steps of the apperceiving self from within the first-person perspective (much like Descartes exhorts us to do the same as the Meditator). Now, of course, the problem is that as interpreters we (as well as Kant himself) need to put our theoretical thoughts *about* thinking into language, which might seem to create the (potential) level confusion that Stephenson describes.

This brings me to my second point in reply to his critique of a level confusion. Stephenson himself conflates language and thought. Contrary to what Stephenson's caricature of it ("Newton's *Principia* displays a Latin syntax; objects in motion do not" [2014:83]) suggests, my derivation of the categories is not a piece of mere language analysis: it is evidently not because the *term* 'necessity' (or indeed 'existence') appears in the sentence that it pertains to thought. The sentence at issue is: "Necessarily, if the 'I think' exists, then the rule that all my representations are accompanied by the 'I think' is satisfied" (premise D7 in the schematisation of the argument; KDA, p. 123).³⁰ But this sentence is just a condensed form of the more extensive argument that the modal category of necessity is an integral element of discursive thought. However, as much as the term 'I think' used in the above-quoted sentence is not to be interpreted as if the term pertained to the identity of discursive thought and could a fortiori be derived from it (well, yes, of course 'I think' belongs to discursive thought, because it *is* that; but it is not a category!), the *term* 'necessity' should not be seen as the ground for deriving necessity from apperception.

When, in his formulation of the third postulate, Kant writes that necessity applies to the thought of the object of experience because that "whose connection with the actual is determined in accordance with

general conditions of experience is (exists) *necessarily*" (A218/B266), I take it that Stephenson would not similarly charge Kant with fallaciously inferring from the appearance of the term 'necessarily' in the above sentence that the category of necessity applies to objective experience. Kant just explains here the conditional necessity that lies in the claim that if anything is actually sensed in accordance with the material conditions of experience, and it agrees with the formal enabling conditions of experience, then it exists; it is under these conditions that the category of necessity applies to the object of experience. Likewise, my sentence D7 just *expresses*, in perhaps irresponsibly condensed form, thought itself and its constitutive modal characteristics (possibility [*if... then*], actuality [*exists*], and necessity [*necessarily*]). In fact, read carefully, the sentence encapsulates the conditional necessity that is expressed by the analytic principle of apperception ("The I think *must be able* to accompany all my representations") rather well.

Stephenson concludes his exposé by alleging that, even if the worry about the level confusion described above could be allayed, my deduction would still not prove that the categories are applied "*to the objects of experience*", but at most that they are "*instantiated in the experience of objects*" (2014:84). But again, it seems to me that Stephenson misguidedly introduces a distinction here that is not pertinent to TD: for if (a) the categories pertain to the unity of apperception, (b) the unity of apperception *is* an objective unity of apperception (B137), and, as per the arguments of B138, A111 and B197/A158 (and of course the full monty of the 'second step' of the B-Deduction), (c) the objective unity of apperception is a necessary and (formally) sufficient condition of both the experience of the object and the object of experience, then (d) the categories apply *to* the objects of experience (QED). Stephenson might object to premise (c) here, as we saw earlier in Sect. 2.6 and briefly again below in the conclusion to the current section—but it is clearly Kant's view. In Kant's theory of *possible* experience, there is no discrepancy between the instantiation of the categories *in* experience and the application of the categories *to* the objects of experience. And as I argued in KDA, the conclusion that the categories apply *to* the objects of experience can, by virtue of a step-by-step progressive derivation of the categories, be inferred from the very principle of thought, the

analytic principle of the apperception (a short version of the argument will be rehearsed below in Chap. 4).

In his critique of my stance, Stephenson shows himself to be a plain old realist, who assumes the existence of objects as distinct from, and outside of, our experience of them. For example, towards the end of his essay, Stephenson claims that one cannot argue that the derivation of the category of 'cause-effect' from apperception is in fact related to cause-effect relations *between objects* (in the real world). That "the categories must apply *to the self and its manifold* if it is to have objects, i.e., representations with objective purport, experience [...] says nothing about our categorization *of objects*", Stephenson (2014:84) reasons. In other words, the "connection" that can be claimed to exist between the "synthetic unity of apperception and its manifold [...] is not a connection that holds between *distinct, conditioned, spatiotemporal substances*" (Stephenson 2014:84).³¹ But when we read e.g. in the A-Deduction that the "understanding [...] is itself the legislation for nature, i.e., without understanding there would not be any nature at all, i.e., synthetic unity of the manifold of appearances in accordance with rules; for appearances, as such, cannot occur outside us, but exist only in our sensibility" (A126–7) (see further Chap. 7), then Stephenson's assertion of a disconnect between, on the one hand, the synthetic unity of the understanding that is required to *experience* connections between objects and, on the other, any causal connection between the objects themselves, or more precisely his denial that synthetic unity in the manifold *is* the connection among objects, seems flatly contradicted by Kant. It is Stephenson's own realist stance that prevents him from understanding the nub of Kant's reciprocity thesis.

2.8 Can 'Contingency' Really Be Deduced from Apperception?

I am delighted that Quarfood finds much to agree with in KDA and that, unlike Dyck and Stephenson, he is on board with the general idea of an a priori derivation of the categories from the principle

of apperception, even though he believes it might still reasonably be argued that the precise manner of such a derivation is ultimately “closed for our investigation” (Quarfood 2014:88), or, that such a derivation is reconstructed in a way different from the one I put forward in KDA. However, Quarfood raises an important objection to my reading of how the negative counterpart³² of necessity, namely contingency, and concomitantly perhaps also the negative counterparts of the other modal categories possibility (i.e. impossibility) and existence (i.e. non-existence) can be derived from the ‘I think’.

The problem he signals concerns the issue whether in the derivation of the category of contingency (and *mutatis mutandis* impossibility and non-existence) the reference “to representations not belonging to those that I apperceive (*my* representations) is legitimate”, given that, as Quarfood correctly observes, the “idea of a derivation from apperception is closely related to the idea of a first-person perspective” (2014:90), from within which the derivation should take place. So the natural question arises: from within such a first-person perspective, can one in fact describe an essential characteristic of representations that are by implication not effectively accompanied by a self-conscious ‘I’? Representations that are not effectively accompanied by an ‘I think’ (in KDA, I called these P3- and P4-representations, the former of which are potentially taken up by apperception and the latter of which are representations that are barred from being so accompanied) are representations that are neither necessarily related, nor accessible, to the ‘I think’, but are just represented by some representer R. Thus it seems that any attempt to describe the nature of such unaccompanied representations, representations that *I* do not presently accompany and so am not aware of existing, is purely speculative and, if at all possible, could only take place from a “third-person (external or metaphysical) perspective” (Quarfood 2014:91), that is, from the perspective of R.

As Quarfood rightly notes, “in conceiving of the non-apperceiving representer R, we are not reenacting the cognitive steps of a judging self” (2014:91), so that it is debatable whether on the basis of the inferred possibility of having P3- or P4-representations the category of contingency can be deduced. So the dilemma appears to be: either

we proceed consistently from the first-person perspective, but then we cannot deduce contingency and a fortiori cannot determine the contingent nature of P3/P4-representations, or we attempt to determine the contingent nature of P3/P4-representations by some external route and thus get a grasp of contingency, which means to give up on the strict derivation *from apperception*.

I am not sure whether I have an adequate answer to this dilemma, but here goes. In KDA (p. 106ff.), I argued that the necessary possibility of P3- and P4-representations can be inferred logically from the principle of apperception, so that the negative counterparts of the first two modal categories, namely impossibility and non-existence, are logically entailed by the positive modal categories of possibility and existence. Put succinctly, if it is necessarily possible that I accompany all and only my representations, whereby the indexicals 'I' and 'my' must be regarded as rigid designators, then it is logically implied that it is necessarily possible that representations are represented (in my head) that are not actually, or even potentially, accompanied by my thinking 'I'. By implication, I have no access to these representations, for I only ever think *my own* representations (which I called P1-representations).

In KDA, I argued that P3-representations might or might not be accompanied in future, and that P4-representations are such that they are not even potentially accompanied. I should clarify that P3-representations are not to be seen as occurrent representations, unaccompanied by an 'I think', that are then potentially later apprehended by an 'I think'. It now strikes me as too vague to speculate about a representation's mere disposition to being apperceived, not least because of the fact that once a representation is accompanied by an actual 'I think', and so is a P1-representation, it can no longer be considered a P3-representation (more precisely, it could not ever have been a P3 if what I can ever think are only *my own* representations). For that reason, I no longer hold that P3-representations are even formally distinguishable from P4-representations. Either representations are (P1) or are not (P3/4) accompanied by an 'I think', which exhausts the logically possible implications of the 'I think'-proposition at B131. The necessary

possibility that is implied in the principle of apperception concerns the fact that I should be able to *thinkingly* accompany all of my own representations without being psychologically forced to always so accompany the *representing* (any representing) that is going on in my head.

But what bearing does this have on the negative counterpart of the third modal category, i.e. contingency? Can it be derived from the principle of apperception? As Quarfood points out, in KDA I concluded that contingency concerns what is implied by the conditional that “necessarily, if the ‘I think’ exists, then the rule that all my representations are accompanied by the ‘I think’ is satisfied”, namely the implication that “[i]t is contingent that for representations for which the rule that all my representations are accompanied by the ‘I think’ is not satisfied, the ‘I think’ is in fact instantiated” (KDA, p. 123). In other words, contingency would have to do with the fact that for all those representations for which the conditions under which they would be apperceived by an ‘I think’ are not satisfied, it is not necessary that they are accompanied by an ‘I think’ in order to be represented.³³ These representations are merely subjectively valid, in the sense of being nothing *for* the ‘I’ of transcendental self-consciousness (which does not necessarily mean that those representations are not consciously apprehended in some subcognitive sense). Of course, from within the first-person perspective one could not have an explicit (reflective) awareness of these merely subjectively valid representations, since the very moment one would have such a reflective awareness one would accompany them with an ‘I think’. Contingency, then, is just the negation of the necessity—which is a conditional necessity—that is expressed by the principle of apperception. Contingency is the characteristic of all representations that are not accompanied by the ‘I think’ of transcendental apperception, which first bestows necessity on representations, insofar as they have been united by this act of apperception.

In later sections of the B-Deduction, Kant characterises this contingent aspect in terms of the merely subjectively valid relation of representations in empirical apperception in contrast to the objectively valid relation of representations, which constitutes a necessary relation between representations, as is quintessentially expressed in a judgement (see in particular §18, B139–40, and §19, B141–2). The contingency

of P3/P4-representations concerns the way they are contingently related in an empirical unity of the mind, as they are prompted in accordance with empirical laws of association or in the manner in which a mind is neurologically wired. Again, here the specifically contingent nature of the way these representations are related can only be *inferred* from within the first-person perspective of the derivation of the modal categories, in the same way that their non-existence before the 'I' can only be inferred as a necessary possibility from the principle of apperception: from within the perspective of the derivation, contingency is the negation of what is objectively valid and necessarily related and as such determinable in conformity with the rules for such determination (i.e. the categories as a whole). From within the perspective of apperception, contingency is thus only determinable as a negative concept. We can of course not determine the very contingent, i.e. *merely* subjective, nature of representations themselves. For we have no metaphysical insight into the absolute constitution of our mind in terms of what really goes on at the subcognitive, subjective level, outside of our self-conscious, transcendental perspective.

I am not sure whether Quarfood's (2014:91) suggestion of looking to the categories of quality will greatly help in deriving the category of contingency, any more than just focusing on the modal categories themselves would. For here too, the intensive magnitude of a sensation, which any sensation, as the matter of a representation, must have, can only be anticipated to the extent that a sensation must be determined to have *some* degree of intensity as a mapping onto the interval $[0,1]$. As Quarfood is right to stress, it is wholly contingent, and so from an a priori (transcendental) perspective entirely unpredictable, *which* degree any arbitrary sensation happens to have. But I do not think that even if via this route the concept of contingency could be derived, it could be done in a way that would circumvent the conundrum that we encountered above with the modal categories, for any determination of the *specific* degree of intensity of some sensation x on the interval $[0,1]$ could similarly only occur via third-person routes (by means of a neurophysiological diagram, say), as indeed Quarfood himself admits: "[I]t too presupposes knowledge that would be transcendent for the apperceptive consciousness" (2014:92).

I think what shows up here are the limits of the Kantian transcendental perspective, or indeed the limits of transcendental self-consciousness as the a priori form of discursive knowledge. On such a notion of self-consciousness as constitutive of *objective* knowledge, we are not transparent to ourselves, obviously not with regard to our putative noumenal selves, but nor even with regard to our phenomenal, empirically contingent, selves (cf. B157), since our very reflexive type of self-consciousness bars us from gaining knowledge of what goes on in the mind on a subcognitive, non-objective level (i.e. in terms of occurrent representations that are not accompanied by an 'I think'). Ironically, on account of Kant's radical subjectivism about the possibility of knowledge we are thus at the same time barred from accessing what is truly *merely* subjectively valid, that is, our most intimate psychological experience, or self-consciousness as more commonly understood. This shows that Kant's radical subjectivism is not a psychological subjectivism. In this context, it seems quite reasonable that Kant's earliest critics, Fichte in particular, found fault with his account—or more precisely the lack thereof—of the sui generis nature of self-consciousness.³⁴ But I think that the general prospects for a theory that purports to give us an insight into the subcognitive terrain of the mind *from the perspective of self-consciousness* are rather slim.

Notes

1. More precisely, Reich is concerned with deriving the *functions of judgements* from the objective unity of apperception, for which he looks for textual support outside the *Critique*. In KDA, I claimed that the *categories* are derivable from the unity of apperception (the 'I think'), but since the categories *are* the functions of judgement, insofar as they determine intuitions as objects, de facto my claim comes down to the same as Reich's. Unlike Reich, however, I contend—and this was my novel claim—that the evidence for the derivation claim can be gathered from the arguments in TD itself.
2. Well, of course Hegel is one who disputes that discursivity is the most basic fact about our thought. See KDA, pp. 8–12 and Chap. 8 (this volume).

3. On the constitutive features of discursivity, see further Chap. 3 (this volume), and KDA, Chap. 5.
4. In the ‘second step’ of TD, which was not discussed in KDA, Kant further argues for the necessary connection between the *thought* of an object and the *perception* of an object; I expound on issues relating to the ‘second step’ in Chap. 7 (this volume).
5. But even here, there is an ambiguity: does that more modest argument mean that the categories are necessarily applicable to experience of objects only, or also to the objects of experience? For an account of these issues, see further Chap. 4 (this volume).
6. That is to say, often one complains that in TD Kant is not specific enough about what the categories are and how they are supposed to be applied to experience, or that the deduction of the categories is not “complete” until the Analogies, or even until the *Metaphysical Foundations of Nature*. Hence, commentators turn to the Analogies, because only there does Kant, so they argue, enter into detail about the particular categories and their application to experience. But this is to confuse the roles of the Analogies and TD. I agree with Michael Friedman that the *Critique* and the *Metaphysical Foundations* “have different yet complementary perspectives on [the] same phenomenal world, about which they establish different yet complementary conclusions” (2015:563–564). I disagree with Friedman’s overall stance though that the scientific laws addressed in the *Metaphysical Foundations* are entailed by the transcendental principles of experience addressed in the *Critique*, such that the superseding of those empirical laws post-Einstein has, as Friedman argues, a direct bearing on the status of the transcendental principles of experience.
7. In KDA, I concentrated on the B-Deduction, but the argument applies, mutatis mutandis, to the A-Deduction as well.
8. See Dyck (2014), Quarfood (2014) and Stephenson (2014).
9. The passage at B127–8 seems less clear-cut in my opinion, as Kant here refers to the kind of derivation of the pure concepts of the understanding (or ideas, in their case) that Locke and Hume had in mind, that is, “an *empirical* derivation” or a derivation from experience, namely in accordance with the psychological laws of association. This would appear to be an inductive derivation from experience as a premise, that is, one that is mutatis mutandis comparable to Kant’s deductive derivation from thought itself (Kant speaks of Locke’s “physiological derivation” [A86/B119], but also in terms of an “empirical deduction”

[B117/A85]; see further KDA, Chap. 3). In both the Lockean and Kantian cases, some sort of logical inference or reasoning is at work, albeit that in the one (Locke's) case the inference is from a psychological principle or empirical fact(s), and in the latter (Kant's) case it is from a general principle (or logical facts) of thought.

10. Cf. Kant's letter of 7 August 1783 to Garve (Br, 10:340), which I quoted in KDA, p. 218n.1.
11. There is some *prima facie* support for Dyck's stance. Recently, Aportone (2009:169–171) has equally argued that indeed the functions of the understanding are just given, “and so discovered, not derived or produced”. Aportone points to A70/B95, where Kant writes: “If we abstract from all content of a judgment in general, and attend only to the mere form of the understanding in it, we *find* that the function of thinking in that can be brought under four titles, each of which contains under itself three moments” (emphasis added). I do not think though that Kant's use of *finden* here is meant to exclude the logical derivability of the functions of thought, and instead proves that Kant thinks we merely come across them. Rather, I take the construction “find that...” to indicate a logical, ostensive procedure of *seeing* that if we do *x*, then *y* follows, or at least to indicate a sufficiently justified claim. Thanks to Wolfgang Ertl and Robert Hanna for discussion.
12. In KDA, p. 10, I point out that it has of course been argued with reference to B145–6, by Krüger (1968), that Kant did not intend to derive the categories, or the functions of judgement for that matter, from apperception, but rested content with their *de facto* status, which appears to be what Kant suggests at B145–6, where he states that “a further ground” for the fact that the unity of apperception consists of “precisely this kind and number” of categories cannot be offered. But compare Wolff's (1995:180–181) reply to Krüger, in which he points out, correctly, that no *further* ground can be given, which does not of course mean *no* ground is provided; and in fact, this remark at B145–6 occurs in the first paragraph after the first part of the deduction's account of how the categories “arise [...] merely in the understanding” (B144) has been completed. Krüger's point is raised again by Quarfood in his critique while suggesting it might still be a viable rival of my interpretation to argue “that the way [the categories] spring from apperception is impossible to elucidate, so that we must stop before the fact that they do, without demanding further explanation” (Quarfood 2014:87–88).

13. Even Locke and Hume believe that the ideas derive from certain dispositions of the mind, i.e. the way sensations are worked up into more complex representations, for which they present meticulous arguments about how this supposedly works, even if they do not believe that this can be shown in any purely a priori manner comparable to Kant's method of proof. See again note 9 above.
14. However, Henrich is rather elusive about what exactly the relation between the categories and self-consciousness consists in. See e.g. Henrich (1976:89).
15. See e.g. Proops (2003:223–224). On Proops's interpretation, see further KDA, pp. 41–43.
16. Dyck refers mistakenly to §§8–10.
17. At B133, Kant writes: "Therefore it is only because I can combine a manifold of given representations *in one consciousness* that it is possible for me to represent the *identity of the consciousness in these representations* itself, i.e., the *analytical* unity of apperception is only possible under the presupposition of some *synthetic* one."
18. There is a similar view in especially the German literature that says that the analytic unity of the 'I think' in relation to the manifold representations that it accompanies is not yet the consciousness of the *identity* of the 'I'. See e.g. Henrich (1976) and Cramer (1990:173–175). But the mistaken assumption here, explicitly so in Cramer, is that the 'I think' at B131–2, which accompanies the manifold jointly (*insgesamt* is the term Kant uses in these passages), is the same empirical consciousness that accompanies *each* single representation in the manifold, of which Kant speaks at B133 and about which he says that it does not thereby constitute a relation to the identity of the subject. But the two types of accompaniment are not the same. That this is not the case is argued in KDA, pp. 176–181.
19. Here, Kant in fact identifies the identity of the self as an analytic unity of apperception of representations.
20. Cf. Prol §47, 4:334: "This thinking self (the soul), as the ultimate subject of thinking, which cannot itself be represented as the predicate of another thing [cf. the same reasoning in B132], *may now indeed be called substance*" (Kant 2002:126; emphasis added); A400: "Now mere apperception ('I') is substance in concept, simple in concept, etc., and thus all these psychological theorems are indisputably correct."
21. See e.g. at B401/A343: "[T]he mere apperception 'I think', which even makes all transcendental concepts possible, which say 'I think

substance, cause, etc.' [...].” I am of course aware of the passage at B406–7, which Dyck quotes as evidence against my view; I refer to it in KDA, p. 94.

22. After all, the concept of subject (*subiectum*) is the translation of the Greek ὑποκείμενον, which in Latin is also translated as *substantia*. Aristotle defines ὑποκείμενον as “that of which the rest are predicated, while it is not itself predicated of anything else [κατ’ οὗ τὰ ἄλλα λέγεται ἐκείνο δὲ αὐτο μηκέτι κατ’ ἄλλου]” (*Metaphysics*, 1028b 36). Compare this with Kant’s definition of the ‘I’ as “the subject, in which thoughts inhere only as determinations, and [which] cannot be used as the determination of another thing”, which he then associates with the concept of substance (A349); and see also of course the description of self-consciousness or the representation ‘I think’ at B132 as that “which must be able to accompany all others [i.e., all other representations] and which in all consciousness is one and the same, [and] cannot be accompanied by any further representation”, in other words, as the ultimate subject. Cf. La Rocca (2003:28–29). La Rocca traces the development of Kant’s view on the relation between substance and the subject or the ‘I’. Notice that, as La Rocca points out, already for the medievals, *subiectum* and *substantia* can “coincide”, but need not be “equivalent”: every *substantia* is a *subiectum*, but not every *subiectum* is a *substantia*.
23. See Schulting (2016).
24. See again La Rocca (2003) for an important account of the grounds for Kant’s abandoning of the idea that the subject *is* a substance, which does not focus on the role of the paralogisms.
25. And even there, Kant tries to “swat away”, as Pippin (2005:211n.6) aptly puts it, what can only be seen as a “garden-variety” modern scepticism. The sceptical challenge that Kant is at most interested in is the Humean denial that our pure concepts are objectively valid in the sense that they tell us the truth about objects.
26. Not Strawson (1968), Stroud (1968) and Rorty (1970), mind. They are participants in the transcendental argument debate that Stephenson is referring to, not the debate about the presumed gap that I discuss in KDA.
27. See Stroud (1968).
28. In KDA, my account of judgement is underdeveloped; I only pointed out how the definition of judgement is the direct corollary of the

reciprocity thesis (cf. KDA, Sect. 10.2). But see further Chap. 3 (this volume), where I elaborate on the theme of the objective validity of a judgement.

29. Some point to Kant's term *Wissen* (B850ff.) as the ostensible Kantian equivalent of 'knowledge', as distinguished from *Erkenntnis* as the ostensible Kantian equivalent of 'cognition'. However, Kant's account of *Wissen* at A822/B850 cannot be read in such a way that the transcendental conditions for *Erkenntnis* are not sufficient for *Wissen*, which Kant defines as a "taking something to be true" that is "both subjectively and objectively sufficient". Transcendental apperception as the transcendental ground of *Erkenntnis* and of truth, as argued in TD, is precisely concerned with a "taking something to be true" that is "both subjectively and objectively sufficient" (cf. A125–7). There is nothing beyond what is known in terms of *Erkenntnis*, that would first be satisfied by *Wissen*. It is simply an anachronism to map a contemporary distinction, which is moreover informed by an Anglophone analytic tradition that does not appreciate the Kantian distinction between the transcendental and the empirical, onto an ostensible distinction between Kant's *Erkenntnis* and *Wissen*.
30. Stephenson goes into exaggeration mode when he flipantly suggests that the concepts 'rule', 'representation', 'accompaniment', or 'satisfaction' could likewise be derived, just because they appear in the language employed to expound Kant's theory.
31. It should be noted that Kant does of course affirm the existence of *things in themselves* as distinct from our minds, but not their quality as appearances, namely, as material objects, bodies, which is what Stephenson here assumes is the case. For further discussion, see Chap. 4, Sect. 4.10.
32. In KDA, my use of the term 'contrary' was misleading, as Quarfoot (2014:89n.2) rightly points out. It should strictly speaking be 'contradictory', but I prefer to use the term 'negative counterpart' to suggest the complementarity between the two contradictories of each modal moment.
33. To which I would now add that, strictly speaking, it is not even *possible* for these representations to be accompanied by an 'I think', given that the conditions under which 'I think'-accompaniment takes place are not satisfied.
34. See Schulting (2017).

References

- Allison, H. 1996. *Idealism and Freedom. Essays on Kant's Theoretical and Practical Philosophy*. Cambridge: Cambridge University Press.
- Ameriks, K. 1978. Kant's Transcendental Deduction as a Regressive Argument. *Kant-Studien* 69 (3): 273–287.
- Ameriks, K. 2003. *Interpreting Kant's Critiques*. Oxford: Clarendon Press.
- Ameriks, K. 2005. A Commonsense Kant? *Proceedings and Addresses of the American Philosophical Association* 79 (2): 19–45.
- Aportone, A. 2009. *Gestalten der transzendentalen Einheit. Bedingungen der Synthesis bei Kant*. Berlin and New York: de Gruyter.
- Callanan, J. 2011. Normativity and the Acquisition of the Categories. *Bulletin of the Hegel Society of Great Britain* 63: 1–26.
- Carl, W. 1989a. *Der schweigende Kant*. Göttingen: Vandenhoeck & Ruprecht.
- Carl, W. 1989b. Kant's First Drafts of the Deduction of the Categories. In *Kant's Transcendental Deductions: The Three Critiques and the Opus postumum*, ed. E. Förster, 3–20. Stanford: Stanford University Press.
- Cramer, K. 1990. Über Kants Satz: Das: Ich denke, muß alle meine Vorstellungen begleiten können. In *Theorie der Subjektivität*, ed. K. Cramer et al., 167–202. Frankfurt a/M: Suhrkamp.
- Dyck, C. 2014. The Function of Derivation and the Derivation of Functions: A Review of Schulting's *Kant's Deduction and Apperception*. *Studi kantiani* XXVII: 69–75.
- Friedman, M. 2015. *Kant's Construction of Nature. A Reading of the 'Metaphysical Foundations of Natural Science'*. Cambridge: Cambridge University Press.
- Guyer, P. 1992. The Transcendental Deduction of the Categories. In *The Cambridge Companion to Kant*, ed. P. Guyer, 123–160. Cambridge: Cambridge University Press.
- Henrich, D. 1976. *Identität und Objektivität. Eine Untersuchung über Kants transzendente Deduktion*. Heidelberg: Carl Winter.
- Henrich, D. 1989. Kant's Notion of a Deduction and the Methodological Background of the First Critique. In *Kant's Transcendental Deductions: The Three Critiques and the Opus postumum*, ed. E. Förster, 29–46. Stanford: Stanford University Press.
- Kant, I. 1992. *Lectures on Logic*, trans. and ed. J.M. Young. Cambridge: Cambridge University Press.
- Kant, I. 2001. *Lectures on Metaphysics*, trans. and ed. K. Ameriks and S. Naragon. Cambridge: Cambridge University Press.

- Kant, I. 2002. *Theoretical Philosophy after 1781*, trans. and ed. H. Allison et al. Cambridge: Cambridge University Press.
- Kant, I. 2005. *Notes and Fragments*, trans. and ed. P. Guyer et al. Cambridge: Cambridge University Press.
- Krüger, L. 1968. Wollte Kant die Vollständigkeit seiner Urteilstafel beweisen? *Kant-Studien* 59: 333–356.
- La Rocca, C. 2003. *Soggetto e mondo. Studi su Kant*. Venice: Marsilio.
- Longuenesse, B. 1998. *Kant and the Capacity to Judge*. Princeton: Princeton University Press.
- Longuenesse, B. 2006. Kant on A Priori Concepts: The Metaphysical Deduction of the Categories. In *The Cambridge Companion to Kant and Modern Philosophy*, ed. P. Guyer, 129–168. Cambridge: Cambridge University Press.
- Mohr, G. 1991. *Das sinnliche Ich. Innerer Sinn und Bewußtsein bei Kant*. Würzburg: Königshausen & Neumann.
- Pereboom, D. 2001. Assessing Kant's Master Argument. *Kantian Review* 5: 90–102.
- Pippin, R. 2005. *The Persistence of Subjectivity. On the Kantian Aftermath*. Cambridge: Cambridge University Press.
- Proops, I. 2003. Kant's Legal Metaphor and the Nature of a Deduction. *Journal of the History of Philosophy* 41 (2): 209–229.
- Quarfood, M. 2014. A Note on Schulting's Derivation of Contingency. *Studi kantiani* XXVII: 87–93.
- Reich, K. 2001. *Gesammelte Schriften*, ed. M. Baum. Hamburg: Meiner.
- Rorty, R. 1970. Strawson's Objectivity Argument. *Review of Metaphysics* 24 (2): 207–244.
- Schulting, D. 2012. *Kant's Deduction and Apperception. Explaining the Categories*. Basingstoke and New York: Palgrave Macmillan.
- Schulting, D. 2016. Review of C. Dyck, *Kant and Rational Psychology* (Oxford UP 2014). *Studi kantiani* XXIX: 185–191.
- Schulting, D. 2017. Apperception, Self-Consciousness, and Self-Knowledge in Kant. In *The Palgrave Kant Handbook*, ed. M. Altman. London and New York: Palgrave Macmillan.
- Seeberg, U. 2006. *Ursprung, Umfang und Grenzen der Erkenntnis. Eine Untersuchung zu Kants transzendentaler Deduktion der Kategorien*. Hamburg: Philo.
- Stephenson, A. 2014. A Deduction from Apperception? *Studi kantiani* XXVII: 77–85.
- Strawson, P.F. 1968. *The Bounds of Sense*, second printing. London: Methuen.

Stroud, B. 1968. Transcendental Arguments. *Journal of Philosophy* 65 (9): 241–256.

Wolff, M. 1995. *Die Vollständigkeit der Kantischen Urteilstafel*. Frankfurt a/M: Klostermann.

Kant's Radical Subjectivism

Perspectives on the Transcendental Deduction

Schulking, D.

2017, XVI, 442 p., Hardcover

ISBN: 978-3-319-43876-4