

## A DEDUCTION FROM APPERCEPTION?

Andrew Stephenson  
Trinity College, University of Oxford

March 2014

### Abstract

I discuss three elements of Dennis Schulting's new book on the transcendental deduction of the pure concepts of the understanding, or categories. First, that Schulting gives a detailed account of the role of each individual category. Second, Schulting's insistence that the categories nevertheless apply 'en bloc'. Third, Schulting's defence of Kant's so-called reciprocity thesis that subjective unity of consciousness and objectivity in the sense of cognition's objective *purport* are necessary conditions for the possibility of *one another*. I endorse these fascinating but unfashionable claims and sketch my own version of what they amount to, which is quite different to Schulting's own construal. I point to some fundamental limitations and problems for Schulting's position and argue that his project needs to be reshaped or at least reconceived in the face of them. Even if Schulting's argument is sound, it does not provide a *deduction*, properly speaking, of the categories.

Dennis Schulting has written a highly original book on that most scrutinized and controversial of philosophical arguments, Immanuel Kant's Transcendental Deduction of the Pure Concepts of the Understanding. The book is tightly focused, tightly argued, and although it is often difficult it is also admirably clear. Schulting claims to find in the so-called first step of the Deduction a derivation of each of the twelve categories from a single principle, the formal 'I think' or the pure original unity of apperception.

There is much in Schulting's book worthy of close critical attention but here I want to focus on three aspects I find especially striking. First, its detailed discussion of *all* of the categories. In this Schulting takes his lead from Klaus

Reich and Michael Wolff, who have done the same for the logical functions of judgement.<sup>1</sup> Schulting focuses directly on the categories themselves, and he also locates particular derivations in Kant's own text – for Schulting, unlike for Reich and Wolff, Kant provides a derivation of each and every category in the Deduction itself. Yet, and this is the second point, Schulting insists that the categories apply 'en bloc':

No category comes apart from the other categories, although in any one actual empirical judgement, the surface form of judgement of course always takes a particular judgemental *logical form*... In the whole of possible experience, however, *all* categories are necessarily instantiated for experience, any instance of experience, indeed to *be possible*.<sup>2</sup>

Third and finally, Schulting mounts a sustained defence of what Henry Allison has called Kant's reciprocity claim, that «unity of consciousness is necessary *and sufficient* for objectivity» (my italics).<sup>3</sup> The standard view on this thesis is that, at best, the conditional can be run in one direct but not the other. Despite Kant's efforts there remains a gap, a certain looseness of fit, between the subject and its object. Schulting argues that properly understood Kant's commitment to full equivalence is in fact warranted.

I agree wholeheartedly with each of these rarely made claims. At least in spirit. As we shall see, I am not sure I can agree with all of the details in how the claims are cashed out. Schulting orients his discussion around a novel take on an established methodological division and it will be useful to begin here.

A 'progressive' transcendental argument is one that does work against the sceptic. Its basic premise is one shared by the sceptic. It then attempts to show that a necessary condition of this premise is something else that the sceptic purports to reject. A 'regressive' transcendental argument is not aimed at the sceptic at all. Its basic premise is not something that the sceptic would accept, so although it follows the same procedure of uncovering necessary conditions, it would be dialectically useless against her. Nevertheless, uncovering necessary conditions can be interesting in its own right and the motivation for regressive

<sup>1</sup> K. Reich, *Die Vollständigkeit der Kantischen Urteilstafel*, Hamburg, Meiner, 1986; M. Wolff, *Die Vollständigkeit der Kantischen Urteilstafel*, Frankfurt, Klostermann, 1995.

<sup>2</sup> D. Schulting, *Kant's Deduction and Apperception. Explaining the Categories*, Palgrave Macmillan, Basingstoke, 2013, p. 12.

<sup>3</sup> Ivi, p. 53.

transcendental arguments can usually be expressed in terms of validation, explanation, revelation, or some combination of these.<sup>1</sup>

For example, a progressive transcendental argument might attempt to *refute* the sceptic who denies the existence of the external world by arguing that such is a necessary condition for the possibility of our inner states displaying any determinate temporal order. A regressive transcendental argument on the other hand might attempt to *explain* our knowledge of the external world – the existence of which it takes for granted – by *revealing* that it involves as a necessary condition the spontaneous synthesis of an inchoate mass of sensibilia into a comprehensible, time-determined manifold of representations of independently persisting spatiotemporal substances that interact with one another and ourselves according to determinate exceptionless causal laws, whereby it would also in a sense *validate* the cognitive apparatus that it has been revealed is required for this synthesis.

There does not seem to be anything unique about the logical form of such arguments. Transcendental arguments are arguments by modus ponens, typically chains thereof. In the atomic case their basic premise is the antecedent, their conclusion the consequent, their auxiliary premise the conditional as a whole. And of course different sceptics are sceptical about different things, so presumably the labels progressive and regressive are not going to apply absolutely but only relative to some particular sceptic. An argument *a* qualifies as progressive in the face of sceptic *s* just in case the following two conditions are met: (i) the basic premise of *a* is accepted by *s* (ii) the conclusion of *a* is rejected by *s*. Otherwise *a* is regressive in the face of *s*.

At least that's the standard understanding of the distinction. Schulting does something quite different with it and it is not always clear how we are to understand the connection. His contention is that Kant's argument in the first step of the Deduction is both regressive and progressive. The progressive side of the argument takes us from unity of consciousness to objective experience while the regressive side takes us from objective experience to the categories. These turn out to map directly onto the two directions of implication in the aforementioned reciprocity thesis. One of Schulting's central concerns is to provide a 'derivation argument' to show that unity of consciousness in the

<sup>1</sup> Q. Cassam, *The Possibility of Knowledge*, Oxford, Oxford University Press, 2007, pp. 67-84.

relevant sense «is constituted by the categories as the complete set of it's a priori functions».<sup>1</sup> That is, the conclusion of the regressive argument just is the premise of the progressive argument and vice versa.

So far, so good. But who is the sceptical target here? I take it that the premise of the progressive side of the argument is supposed to be beyond doubt for any sceptic. For Schulting interprets Kant's first principle concerning the unity of consciousness analytically.<sup>2</sup> The famous claim near the beginning of the Deduction that «the *I think* must *be able* to accompany all my representations» (*KrV*, B131) is to be understood as a conditional involving de dicto necessity: *If* any given representation is going to be something *for* the subject, as opposed to being merely something *in* the subject or indeed elsewhere, and therefore nothing *for* the subject, *then* necessarily it is possible for the subject to ascribe that representation to itself. In other words, Kant is simply pointing out the grammatical connection between the *I think* and *my* representations. The Kantian cogito, at least at the beginning of the argument, is the empty, formal, logical self. (Schulting's primary target here is Paul Guyer, who has argued that the apperception principle must be understood as synthetic from the outset.<sup>3</sup> For Schulting, analytic unity comes first but can then be shown to be 'rigorously coextensive' with synthetic unity.<sup>4</sup>) We can assume, then, that the first condition of progressivity – condition (i) above – is satisfied for all values of *s*. To determine the particular sceptical target of the progressive side of Schulting's argument we must look to the conclusion. What is objective experience?

Schulting puts it in many different ways but it is clear that objective experience should be understood as the mere *claim* to empirical knowledge and not empirical knowledge itself.<sup>5</sup> Objective experience is experience with objective *purport*; it represents objects *as such*, but does *not* necessarily do so veridically. The problem is that it is not clear what sceptic would deny that experience at

<sup>1</sup> Schulting, *op. cit.*, p. 211.

<sup>2</sup> Ivi, pp. 75-82, 95-110.

<sup>3</sup> See especially P. Guyer, *Kant on Apperception and A Priori Synthesis*, *American Philosophical Quarterly*, 17, 3, 1980, pp. 205-212; P. Guyer, *Kant and the Claims of Knowledge*, Cambridge, Cambridge University Press, 1987; and P. Guyer, *The Transcendental Deduction of the Categories in The Cambridge Companion to Kant* Cambridge: Cambridge University Press, 1992, pp. 123-160.

<sup>4</sup> Schulting, *op. cit.*, pp. 110-113.

<sup>5</sup> Ivi, pp. 56-60, 71-73.

least *seems* to present us with an independent world of causally interacting objects.

I can put this problem in a different way. In defending Kant's reciprocity thesis Schulting wants to deny that there is a gap between subjectivity and objectivity. In particular, he wants to defend Kant's inference from the unity of apperceptive consciousness to the possibility of experience or cognition of objects, understood non-factively. Not only could there be no thinking about objects without the possibility of self-consciousness; nor could there be self-consciousness without the possibility of thinking about objects. But is *this* really where the gap is traditionally supposed to appear in Kant's Deduction? Is *this* the inference that Stroud, Rorty, Guyer and Cassam have all so forcefully brought into doubt?<sup>1</sup> It seems to me that something quite different is at issue. Here are four different ways of getting at the gap that I have in mind: (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.<sup>2</sup>

My point here is not to say that Schulting ought to have addressed this issue. My point is to hone in on the target of the Deduction according to his interpretation. For evidently it is not the sceptic about the existence of the external world, or the sceptic who doubts whether any of our beliefs about objects are true or even justified. In the face of such sceptics Schulting's argument remains wholly regressive. There is a gap that Schulting has not bridged or eliminated as much as moved around or put to one side.

<sup>1</sup> See B. Stroud, *Transcendental Arguments*, *Journal of Philosophy*, 65, 1968, pp. 241-256; R. Rorty, *Strawson's Objectivity Argument*, *Review of Metaphysics*, 24, 1970, pp. 7-44; P. Guyer, *Kant and the Claims of Knowledge*, Cambridge, Cambridge University Press, 1987, p.117; Q. Cassam, *Transcendental Arguments, Transcendental Synthesis and Transcendental Idealism*, *The Philosophical Quarterly*, 37, 149, 1987, pp. 375; J. Van Cleve, *Problems from Kant*, Oxford, Oxford University Press, 1999, pp. 73-104; and A. Gomes, *Is Kant's Transcendental Deduction of the Categories Fit for Purpose?*, *Kantian Review*, 15, 2, 2011, pp. 118-137.

<sup>2</sup> One is reminded by this last formulation of John McDowell's distinction between justification and excuse or exculpation in J. McDowell, *Mind and World*, Cambridge M.A., Harvard University Press, 1996, p. 8.

So what *is* Schulting's gap? Who *is* his sceptical target? One natural answer at this stage would seem to be the following. Perhaps Schulting is happy to rest content with showing that we must apply the categories to the objects of our experience, whether or not this application turns out to be justified and to result in true beliefs. Perhaps he aims to show that *this* is a necessary condition for the possibility of self-consciousness. If we are to be able to be apperceptive unities of consciousness, we must, for example, conceive of all events as having causes. On this reading the sceptical target would seem to be Hume, not in his guise as a sceptic about the external world but rather in his guise as a sceptic about causation. Hume naturalises and subjectivizes the notion of causation, ultimately reducing it to something much less than a relation of *genuinely necessary* connection between distinct existences. He may want to retain the words 'necessity' and 'cause', but he and Kant are no longer talking about the same thing, and as Kant means it, Hume would deny that we must conceive of all events as having causes.

This is a very natural answer. It fits with much of what Schulting says, certainly with much of what Kant says, and it is a common enough interpretation. But it is wrong again – this Hume is not Kant's target on Schulting's reading. To see why, it will be useful to give my own gloss on this kind of position, which will also be my own take on the three central claims of Schulting's book that I outlined at the beginning: that each of the individual categories must be accounted for, that they nevertheless come en bloc, and that Kant's reciprocity thesis properly understood is defensible.

The categories articulate the concept of an object *in general* (*KrV*, A93/B126, B128, B158, B159). There are both unschematized and schematized versions of this concept, depending on whether or not it is understood in abstraction from our particular forms of sensibility, space and time. The unschematized concept of an object in general is something like that of a whole with parts that possesses certain qualities, doesn't possess others, and which can stand in some kind of ground-consequence relations to other such things. Many things qualify as objects in this sense, including abstracta and things in themselves. Even discursive intellects that do not share our forms of sensibility would ultimately, as we do, have to think in terms of it (though it is a more difficult issue whether a non-discursive, intuitive intellect would have to, and I suspect Schulting is right when he implies it would not). The *schematized* concept of an object in

general, on the other hand, is more restrictive, though it still allows for various distinctions in kind. Most generally, it is the concept of an appearance – an object of possible human cognition. This would still include both empirical objects and the pure sensible objects of mathematics, however. Empirical objects are those appearances which affect our senses and are thereby objects of possible human *empirical* cognition, or experience. In this application, then, the schematized categories articulate the concept of an *empirical* object in general, which is something like that of a spatiotemporally extended and located, sensible property-possessing, fully causally functioning particular – a material thing.

Empirical objects are spatiotemporal wholes with spatiotemporal parts, thus a plurality considered as a unity, which is to say, a totality. Any given empirical object possesses some sensible properties and does not possess others – it has its reality limited by negation. And each is a substance in which accidents inhere and which stands in reciprocal causal relations to the whole unified community of all other such objects in and through one spacetime (*KrV*, B110-13). In this way, the schematized quantitative, qualitative, and relational categories are mobilized *wholesale*, as *collectively* articulating the concept of an empirical object in general.

There is no reference to the modal categories here because «The categories of modality have this peculiarity: as a determination of the object they do not augment the concept to which they are ascribed in the least, but rather express only the relation to the faculty of cognition» (*KrV*, A219/B266; cf. A74/B99-100). Nevertheless, in this relation as it is manifested in experience, the modal categories are likewise mobilized wholesale. For in experience, individual empirical objects are comprehended as interconnected parts of a unified causal network in which everything that is (empirically) possible is actual and thus (empirically) necessary (*KrV*, A220-35/B267-B287). This unified causal network is the world, in a sense that single object of our single, all-encompassing experience.

Note that although paradigmatic of empirical objects are things like birds, trees, tables and chairs, what makes the schematized categories in this application a collective articulation of the concept of an empirical object *in general* is that in themselves they say nothing further about any particular empirical determination of quantity, quality or relation – size, shape, colour, particular causal power, and

so forth. (And again, note that modality is absent here, as we would now expect, for it does not have particular determinations in this sense). In this way, though schematized to restrict the categories to the cognizable realm of appearances, and then specified further to the *empirically* cognizable realm, the concept of an empirical object *in general* remains pure (or at least a priori – see *KrV*, B2-3). Not every possible object of human experience is a bird or a tree, but every possible object of human experience is a spatiotemporally extended and located, sensible property-possessing, fully causally functioning particular. Thus it is this concept that is necessarily employed when we have a world in view at all, that is necessary for us to be presented with objects *as such*, for our representations to have objective purport, veridically or otherwise.<sup>1</sup>

That is a sketch of the individual functions of the categories along with their essentially collective employment as a necessary condition for the possibility of experience (with experience understood in a conceptually thick sense, but still not in an epistemically thick sense as actual knowledge). Now regarding the reciprocity thesis.

In one direction the relation here is relatively clear. Kant thinks that possible self-consciousness or apperception is a necessary condition for object-consciousness or experience because it is just part of what it is to be an object to be other than the subject, to be «that which is opposed to our cognitions being determined at pleasure or arbitrarily» (*KrV*, A104). For Kant, being conscious of an object in the relevant *cognitive* sense involves not merely being affected by it and being able to respond to this affection in various ways, but also at the same time involves being aware that it is an object and understanding what an object is – the consciousness of objects in the relevant sense is a consciousness of objects *as such*. And this, for Kant, requires mobilizing the concept of such an object. This is a broadly conceptualist assumption that was also operative above. What is crucial now is that it follows that in being so aware one must also mobilize the concept of a subject, since this concept is also part of the concept of an object – among those other marks I mentioned, an object just is something distinct from the subject. And then the thought is that if one mobilizes the concept of a subject in one's object-consciousness, then it is possible to become conscious,

<sup>1</sup> For further elaboration and discussion see A. Stephenson, *Kant on Non-Veridical Experience*, Kant Yearbook, 3, 2011, pp. 1-22 and A. Stephenson, *Kant on the Object-Dependence of Intuition and Hallucination*, Unpublished Manuscript, 2014.



again in the cognitively rich sense, of oneself *as* a subject. And this just is to become self-conscious. In this way the possibility of object-consciousness is a sufficient condition for the possibility of self-consciousness, so possible self-consciousness is a necessary condition for possible object-consciousness.

In the other direction the relation is a little trickier to appreciate and has been even more roundly criticized. I propose that the key to understanding it runs along the following lines. Kant thinks, conversely, that possible object-consciousness is a necessary condition for possible self-consciousness because he understands self-consciousness as *second-order* consciousness. That is to say, self-consciousness, or apperception, for Kant, is consciousness of one's consciousness of objects as such. Let C(O) stand for the consciousness of an object as an object. Then the general conceptualist commitment in the background of the previous points was that, for there to be consciousness in the proper, cognitive sense, where it is understood what the thing inside the brackets is, a concept of that thing is required. And the current point is accordingly that 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. And if possible self-consciousness requires the concept of an object, then in effect it requires the possibility of object-consciousness. Possible object-consciousness in this way is a necessary condition of possible self-consciousness, so the possibility of self-consciousness is a sufficient condition for the possibility of object-consciousness.

So what makes Schulting's account so different from this one? It shares what I have identified as three key characteristics of his account – an account of the individual categories and their collective employment as well as a defence of the reciprocity thesis connecting subjectivity with (non-factive) objectivity. And it also shares the same progressive/regressive structure – it is, for instance, progressive in the face of Hume (in one guise) but regressive in the face of the Cartesian. Well the difference is that Schulting's account does not seem at all to concern *our application* of the categories *to the objects of our experience*. I will explain.

Schulting's derivation argument has twelve stages corresponding to each of the twelve categories. At each stage – that is, for each category – the conclusion of the argument has the following form:

The category of [insert category] pertains to the identity of discursive thought and hence is analytically derivable from it.

So what does ‘pertains to’ mean? It seems to mean at least two things, each of which is problematic. Two examples will suffice. First, necessity. Here is what Schulting says:<sup>1</sup>

The conceiving of the relation between condition and conditioned yields the third category of modality: necessity... if the ‘I think’ ‘exists’, is posited absolutely, or is instantiated, that is, discursive thought is a given fact, then necessarily the ‘I think’ is posited in accordance with the rule that all my possible representations are posited relative to the ‘I think’...

So from the premise:<sup>2</sup>

Necessarily, if the ‘I think’ exists, then the rule that all my representations are accompanied by the ‘I think’ is satisfied.

It follows that:<sup>3</sup>

The category ‘necessity’ pertains to the identity of discursive thought and hence is analytically derivable from it.

Or so the story goes. But this looks like a confusion of levels, akin to failing to keep separate meta-language and object-language. Schulting seems to argue 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. Such a move looks fallacious. Newton’s *Principia* displays a Latin syntax; objects in motion do not. Perhaps it could be argued that the move is in fact valid for the case at hand, but in doing so one would have to be very careful to avoid becoming susceptible to Gaunilo-style parody. The concept of existence is also mobilized in the premise, so could one not also infer that the category ‘existence’ pertains to the identity of discursive thought and hence is analytically derivable from it? Or indeed likewise for any one of ‘The concept ‘rule’...’, ‘The concept ‘representation’...’, ‘The concept ‘accompaniment’...’, ‘The concept ‘satisfaction’...’. And even if this worry can in turn be allayed, there remains the problem that this is not a category being

<sup>1</sup> Schulting, *op. cit.*, p. 122.

<sup>2</sup> Ivi, p. 123.

<sup>3</sup> Ivi, p. 123.

applied *to* the objects of experience at all. At most it is a category being instantiated *in* the experience of objects. This brings me to the second example, causation.

The key claim this time is the following:<sup>1</sup>

Self-consciousness itself thus shows up a relation of cause and effect, where the activity of the self is the cause of determinacy [in the manifold] as effect, and itself does not rest on a more original determination. Self-consciousness is essentially self-activity or spontaneity.

By now we have made the move from analytic unity of apperception to synthetic unity, but things run precisely as before. From the premise:<sup>2</sup>

The subsisting 'I' is the original synthetic unity of apperception, which is that action which is the power of the self-active subject and spontaneously produces a synthetic unity among the manifold of representations.

It is supposed to follow that:<sup>3</sup>

The category 'cause-effect' pertains to the identity of discursive thought and hence is analytically derivable from it.

To my eye it is not obvious that the same level- and parody-issues aren't still lurking in the background. More importantly, however, what we have here is again a category instantiated *in* the experience of objects and not an application of a category *to* the objects of experience. There are at least two problems with this.

First, this method of deriving the category 'cause-effect' in no sense gives us a justification of our practice of applying this category to the objects of experience. One way of providing such justification would be to show that such a category does apply, that the objects of experience do (in fact or as a matter of necessity) instantiate causal relations (in Kant's sense). Schulting has long conceded that this 'strong' conclusion is not his professed goal, and I think this is fair. But the point is that so far we don't even have the weaker conclusion that our practice of applying categories to objects is in some sense validated or legitimated by the

<sup>1</sup> Ivi, p. 137.

<sup>2</sup> Ivi, p. 147.

<sup>3</sup> Ivi, p. 147.

unavoidability of it. For nothing at all has yet been said about this practice. Schulting's claim is simply that the categories must apply *to the self and its manifold* if it is to have objects, i.e. representations with objective purport, experience. This says nothing about our categorization *of objects*.

The second, closely related problem is that it is not clear how these could be related. That is, one might think that the necessary instantiation of certain concepts by the self and its manifold could somehow ground the self's application of the self-same concepts through its manifold to objects. But there seems to be an important equivocation in exactly what concept is being employed here, and so no reason to think that this could be the case. The category 'cause-effect' as we apply it to objects is that of there being a necessary connection between distinct, conditioned, spatiotemporal substances. This is the concept for which Kant wants to provide a transcendental deduction and which Leibniz, Spinoza, Malebranche, Berkeley, and Hume, all in their own very different ways, rejected. The connection between Schulting's synthetic unity of apperception and its manifold is, granted, necessary in the required Kantian sense, but it is not a connection that holds between *distinct, conditioned, spatiotemporal substances*.

It therefore seems to me that there is no sense in which Schulting's Kant has proven the objective-validity of the category of causation (as applied to objects). It has not been shown to apply to objects. Nor has it been shown to be a necessary condition for us to be given objects as such in experience. For *it* has not been deduced from apperception at all.

So where does this leave us? I do think that there is a sense in which the argument on Schulting's interpretation might still count as progressive. For if successful, and granting certain reasonably plausible principles concerning epistemic self-access, it does seem that Schulting's Kant could have shown that there must *be* a priori concepts. Hume, now in his guise as an empiricist – or, as we might say, as a sceptic about a priori concepts – has become the target. This would be a significant enough result, not least because Hume's empiricism (i.e. scepticism about a priori concepts) plays a major role in his scepticism (about causation and the external world). But can it be called a *deduction*? To me it seems not, because the question at issue here is one «which concerns the fact (*quid facti*)» rather than one 'about what is lawful (*quid juris*)» (*KrV*, A84/B116).