

Transcendental Apperception: Consciousness or Self-Consciousness? Comments on Chapter 9 of Patricia Kitcher's *Kant's Thinker*

RALF BUSSE

Johannes Gutenberg-Universität Mainz

Email: rbusse@uni-mainz.de

Abstract

A core thesis of Kitcher's is that thinking about objects requires awareness of necessary connections between one's object-directed representations 'as such' and that this is what Kant means by the transcendental unity of apperception. I argue that Kant's main point is the spontaneity or 'self-made-ness' of combination rather than the requirement of reflexive awareness of combination, that Kitcher provides no plausible account of how recognition of representations 'as such' should be constituted and that in fact Kant himself appears to lack the theoretical resources to clearly distinguish between (first-level) consciousness and self-consciousness or apperception properly so-called.

Keywords: representation, apperception, consciousness, self-consciousness, function, spontaneity, transcendental

Patricia Kitcher's *Kant's Thinker* is an extraordinarily important, clear and rich book on Kant's mature philosophy. In the following I shall focus on some aspects of its central chapter 9 'Arguing for Apperception'. What I particularly like about the book is, first, that while Kitcher goes into many details of Kant's Critical work, her study has an exciting single core thesis. The thesis is that, due to the rational character of human thinking, consciousness of objects and self-consciousness are mutually interdependent. Secondly, the reader will find pearls of insight everywhere. Let me mention only one from chapter 9. Kant's discussion of 'what is meant by the expression "an object of representations"' in the A-deduction of the *Critique of Pure Reason* (CPR, A104–5) has

been misinterpreted as a development of a bundle-of-representations view of objects of human cognition.¹ Within a few lines Kitcher makes it clear that Kant is not so much concerned with the question ‘What is an object?’, but with the question ‘What is required in order for a human cognizer to *refer* given representations to an object?’ Her concise answer hits the nail on the head: Kant’s point is that, in order to refer a manifold of ‘particular representations of sensible properties’ ‘to a common object’, a cognizer needs to have available a ‘rule that enables her to determine which representations could be representations of the same object’ (p. 130).

1. Thinking in vs. Thinking about Representations

Let me introduce my critical remarks with a terminological distinction. According to Kant, every act of thinking is *thinking in representations*: we think about objects by having and using intuitions and concepts. Quite another thing is *thinking about representations*, or so it might seem: here we make our own representations the objects of our thinking. Compare a parallel distinction with regard to language: all speaking and writing is speaking and writing *in* linguistic symbols. Quite another thing is speaking or writing *about* linguistic symbols, as when one uses quotation marks in order to talk about the words quoted.

The thesis by Kitcher that interests me is that, according to Kant, all thinking *in* representations about (outer) objects already involves or requires thinking *about* one’s own representations. According to Kitcher, Kant makes two core points: first, thinking about objects has to be rational thinking; and secondly, a subject’s act of thinking can be rational only if the subject recognizes her would-be object-directed representations ‘as such’ and is aware of necessary connections between those representations. For Kitcher, this awareness of necessary connections between one’s would-be object-directed representations is what Kant calls the ‘transcendental synthetic unity of apperception’.

2. Representation of Combination ‘as such’

My first critical remark is that I sometimes find Kitcher’s hermeneutic moves in favour of her core thesis a bit tendentious. For example, in §15 of the B-edition Kant writes:

We cannot represent anything as combined in the object without ourselves having combined it beforehand; and ... among all representations *combination* is the only one that ... – being an *act*

of the subject's self-activity – can be performed only by the subject himself. (CPR, B130; trans. following Kitcher; my emphasis)

Kitcher comments: 'His claim is that the sort of combination relevant to cognition requires that combination not simply be effected in some manner, but that it is represented as such, as combination' (pp. 142–3; my emphasis).

However, in the passage above Kant is merely stressing the spontaneity or 'self-made-ness' of combination as an act of representing, not a required second-order awareness of that representational act of combination 'as such'. Kant is indeed working here with a distinction between an act of representing, on the one hand, and what is represented, on the other. But the represented is not a first-order act of combination, which is represented by a second-order representing. Rather, he means that one can represent several different parts or aspects *a, b, c ...* of an object as standing in a relation of combination *R* only by actively connecting one's intuitive representation of part *a*, one's representation of part *b*, that of *c*, etc., by a combining relationship *R'*. Thus, when Kant in §15 says 'Combination is the representation of the *synthetic* unity of the manifold', I think he should not be interpreted as meaning that combination is a second-order awareness of a first-order unified act of synthesis. Rather, by combination he here means a first-order representation of a synthetic unity *in the object*.

To consider a favourite example of Kant's, he insists that one can represent different spatial parts of a line as integrated into one object, a determinate line in space, only by actively 'drawing the line', that is, by actively integrating one's intuitive representations of the line's parts into one complex intuition (cf. CPR, §17, B137–8). I represent line segments *a, b, c ...* as forming one line by actively combining given representations of *a*, of *b*, of *c*, etc. To be sure, a few lines further Kant requires 'synthetic unity of consciousness' (B138) in this act of drawing a line. But there is no evidence that he thereby has in mind a reflexive awareness of one's synthesizing act 'as such'. He appears to be mainly stressing the character of the synthesizing act as a unified conscious (but not necessarily *self-conscious*) performance. Another few lines further in §17, this unified activity does not appear to *consist in* the performance of transcendental apperception, but rather to be a mere necessary condition of the always possible explicit apperceptive act expressed by the locution 'I think' (all these different representations) (B138).

3. The Double Scrutinizing

A second point is that Kitcher explicitly assumes a double layer of mental activity in the representation of objects. She assumes two activities of ‘scrutinizing’, one object-directed and the other directed to one’s own representations, and she takes them to be executions of two different innate tendencies of the subject: in addition to the object-directed activity there is

an extra layer of scrutiny. Cognizers not only come with innate tendencies to scrutinize the contents of representations for indications of causal relations; they *also* have *innate tendencies to scrutinize* not the contents of representations, but *representations themselves* to find representations that meet some other principle [namely, what she calls the ‘I-rule’]. (p. 135; my emphasis)

She supports this interpretation by reference to Kant’s reply to Eberhard’s *Streitschrift* of 1790. But it is hard to find there a hint of a twofold set of basic innate tendencies of the understanding. Quite the opposite, with respect to the categories Kant insists that there is ‘*nothing innate save the subjective conditions of the spontaneity of thought* (in conformity with the unity of apperception) (*nichts Angebournes, als die subjektiven Bedingungen der Spontaneität des Denkens (Gemäßheit mit der Einheit der Apperzeption)*)’ (*Streitschrift*, 8: 223, Kant 2002: 313; my emphasis). A few lines later he is concerned with the pure category of substance, which contains ‘nothing more than *the logical function* (*nichts weiter, als die logische Funktion*)’ (my emphasis).

This suggests that Kant only assumes one single set of basic innate tendencies of the understanding, namely, the twelve ‘functions of thinking’ that he lists, qua functions of articulate judgemental thinking, in the table of judgements (*CPR*, §9). And it suggests that for him conformity with the unity of apperception simply consists in spontaneous *a priori* synthesis in accordance with the basic functions of the understanding. To be sure, Kant assumes at least two different ways of executing ‘the same function’ of the understanding (*CPR*, §10, A79/B104–5). But this is not a distinction between an object-directed use of the basic functions of spontaneity, on the one hand, and a representation-directed use, on the other. It is a distinction between an execution directly at the level of intuitions and one by means of general concepts.

Moreover, Kitcher’s ‘I-rule’ says that representations belong with all others to one consciousness (p. 126). This seems to be a theoretical

statement of Kant's *about* ordinary cognizers, not a rule that we follow explicitly or even implicitly in ordinary cognition.

4. How is Recognition of Representations 'as such' Constituted?

My next point is that it is not very clear how recognition of representations 'as such' is supposed to be constituted. Kitcher writes: 'the unity of apperception is produced through the creation of relations of necessary connection across representational states – and the states thereby come to be recognized as such' (p. 136). Does this not amount to the idea that thinking *about* representations 'as such' results from a certain kind of thinking *in* representations? The idea seems to be that by treating representational states in active thinking as necessarily connected we manage to think about the representational states themselves, rather than merely about their objects or contents. This view strikes me as rather implausible. It would be similar to the claim that by actively inferring 'Socrates is mortal' from 'All men are mortal' and 'Socrates is a man', we arrive at a reference to the three judgements, rather than just to men, mortal beings and Socrates. Yet the inference just seems to be a case of inferential thinking *in* judgements, not one of thinking *about* one's own judgements.

One might object that in an inference one standardly does more than merely state or imply a logical connection between what the premises say and what the conclusion says: one also asserts the premises and the conclusion. But there appear to be ways to cancel these assertions and to solely highlight the logical connection that do not require reference to the sentences or judgements involved. One way is the explicit assertion that from its being the case that all men are mortal and that Socrates is a man it logically follows that Socrates is mortal. Another way to state the necessary connection in a first-order manner might be the strict conditional 'If all men are mortal and Socrates is a man, then, by logical necessity, Socrates is mortal'. Kant holds that a hypothetical judgment (*Satz*) contains a relationship (*Verhältnis*) of two judgements, antecedent and consequent (*CPR*, §9, A73/B98). But by this he hardly means that a hypothetical judgement is a second-order judgement *about* two judgements, any more than a categorical judgement is *about* the subject and the predicate concept, but rather involves them.²

5. Rationality, Spontaneity and Rules

My fourth point is a tentative suggestion. I find Kitcher's observation instructive that Kant is stressing the conditions of empirical thinking being *rational*. But what Kant seems to be mainly insisting on is not the requirement of a special layer of thinking *about* one's own representations

in one's object-directed thinking *in* representations. His core point rather seems to be the spontaneity or self-made-ness of rational thinking and the irreducibility of such acts of thinking to mere collections of receptive and reproductive representations. His main thesis appears like a transcendental-psychologicist version of Kripkenstein's tenet concerning rule-following:³ rational acts such as inferring a conclusion from certain premises are executions of personal competences of rule-following. These competences of rule-following (i) cannot be reduced to the mere having of a heap of mental states (like images) nor to dispositions to behave in certain ways or to be in certain mental states under certain conditions. (ii) Nor can they be explained in terms of explicit awareness of the rules in question, on pain of vicious regress. The central Kantian analogue to the Kripkensteinian irreducible rules would, of course, be the *a priori* functions of the understanding.

6. What is Transcendental Apperception?

Let me finally ask: what is Kant's transcendental apperception? My suspicion is that within Kant's overall theory it is a wooden iron. Kant notoriously oscillates between translating the Latin 'apperceptio' as *Selbstbewusstsein* (self-consciousness) and as *Bewusstsein* (consciousness). For example, he starts the first step of the B-deduction in §16 by introducing the transcendental 'unity of self-consciousness' (B132), but ends up in §20 with highlighting the categories as conditions of the coming together of the manifold of intuition 'in one consciousness' (B143). Self-consciousness, however, is more than mere consciousness. We need to distinguish between at least four kinds of consciousness: (i) purely sensory consciousness (a raw, given sensation of red, or a raw feeling of pain), (ii) first-order intentional consciousness (of particular worldly objects, or of propositions or states of affairs), (iii) second-order, reflexive consciousness of one's own mental states (as in 'That thought a second ago was rubbish, but this thought now is fine'), (iv) second-order, reflexive consciousness of oneself as oneself, or I-consciousness (typically expressed by sentences in the first person singular).⁴

Case (i) is not acknowledged as a kind of consciousness by Kant at all. I doubt that he clearly distinguishes between kinds (ii), (iii) and (iv). More specifically, I doubt that he has the theoretical resources to coherently conceive of versions of second-order consciousness of kinds (iii) and (iv) that deserve to be highlighted as 'transcendental'. In §16 of the B-edition Kant explains that he calls the unity of apperception 'transcendental' 'in order to designate the possibility of *a priori* cognition from it' (B132; italics for bold-face of the Guyer/Wood edition).

This is in accord with his mature use of ‘transcendental’ as meaning *source of synthetic a priori knowledge*. But when Kant stresses the contrast between empirical and transcendental apperception, an earlier use appears to be in play. Kant originally uses ‘transcendental’ as meaning *originating in a purely intellectual manner, that is, purely from the understanding, without any help of either sensation or the pure forms of sensibility*.⁵ Traces of this original use can even be found in later works of Kant’s, for example, in his notion of a ‘transcendental definition’ in the *Critique of Judgement* (XXII, 5: 177, n.).

In fact, Kant seems to be committed to the view that transcendental apperception is a purely intellectual form of I-consciousness. In §16 he writes that ‘[t]he *I think* ... is an act of *spontaneity*, i.e., it cannot be regarded as belonging to sensibility’ (B132). One wonders, however, how a mere act of the understanding should deserve to be classified as consciousness of level (iv), i.e. reflexive consciousness of oneself as oneself that is properly expressed in the first person singular. All reflexive consciousness seems to require some means of forming second-order representations of one’s first-order representations or of one’s first-order thinking activity. But the understanding by itself does not provide any such reflexive means. For Kant, the faculty that first produces a kind of representations of representations (namely, reproductions) is inner sense. Without the help of sensibility, the understanding can at best produce acts of thinking, but not reflexive acts of thinking about one’s own thinking (kind (iii)), not to mention acts of thinking of oneself as a thinking subject (kind (iv)). To be sure, in §§24–5 Kant discusses the interplay of pure understanding and inner sense in the transcendental synthesis of the imagination. He says that it is the activity of the understanding that affects inner sense, just as outer objects affect outer sense. But Kant insists that it is the intellectual part of this interplay that is expressed by the first-person locution ‘*I think*’.⁶

It is no good to simply postulate that the concept *I* or *I think* is *a priori* (as has been suggested to me by Tobias Rosefeldt). Kant rejects innate concepts (see *Streitschrift*, 8: 221, Kant 2002: 312). All that is innate with regard to the understanding are the twelve basic functions of synthesis. However, an execution of these functions at best serves to turn given representations into unified, synthesized representations that refer to some object, i.e. into first-order intentional consciousness (kind (ii) in the classification above).⁷ One would need an account of how the representation *I think* is ‘originally acquired’ (cf. *Streitschrift*, 8: 222, Kant 2002: 313) on the basis of the functions of the understanding.

What is groundbreaking in Kant's view of transcendental apperception is his insight that at the basis of I-consciousness lies, first, thinking activity, and, secondly, an activity of integrating the diverse contents of consciousness. But he does not seem to clearly see that this all-encompassing integrating activity has in turn to be thinkingly taken up in a second-order conceptual representation in order to arrive at an awareness of oneself as a thinking subject. And it is doubtful that his theory of inner sense and of time as its pure form could provide the required reflexive link between first-order thinking activity and the second-order conception of oneself as a thinking subject.

So my worry is that in his notion of transcendental apperception Kant is mixing up mere *a priori*-synthetic thinking consciousness (kind (ii)) with self-consciousness properly so-called (kinds (iii) or (iv)). In the end, I think, Kant is committed to saying that an act of transcendental apperception simply consists in the synthetic, *a priori* activity of thinkingly integrating given representations into the unity of consciousness by the functions of the understanding. But no mere act of thinking, however clear, distinct, encompassing, and *a priori* it may be, is already an act of thinking about one's own thinking, much less one of thinking of oneself as a thinking subject. It seems to be no accident indeed that Kant alternates between translating 'apperception' as 'self-consciousness' or just as 'consciousness'. Thus, while Kitcher tends to find self-consciousness everywhere in rational thinking à la Kant, I find it hard to locate self-consciousness properly so-called anywhere in his theory of the understanding.

Notes

- 1 This interpretation is most explicit in Thöle (1991: 194).
- 2 Sellarsians may hold that both explicit assertions of logical consequences and non-truth-functional conditionals require a subtle meta-linguistic analysis to the effect that sentence tokens of one kind can or ought to be inferred from tokens of another kind (cf. Sellars's (1958) 'inference ticket' view of lawlike statements). But Kant hardly is a proto-Sellarsian in this respect.
- 3 For an introduction to the debate see Miller 2002.
- 4 Here I am taking up Hector-Neri Castañeda's formula 'thinking about oneself qua oneself' for self- or I-consciousness; see e.g. Castañeda 1999: 253.
- 5 See Bencivenga 1987: 3–10.
- 6 See also the delicate pair of claims in the famous footnote of the B-Paralogisms that the 'I think' is an empirical proposition (which hence somehow rests on sensibility), but that the representation 'I' in it is 'purely intellectual' (CPR, B422–3).
- 7 It has been claimed by Brook (1994: 37) and Dickerson (2004: 95) that this is indeed all that Kant has in mind with his notion of transcendental apperception. In his review of Dickerson Rosefeldt (2007) rightly objects that this can hardly be the whole truth, as far as Kant's intentions are concerned.

References

- Bencivenga, Ermanno (1987) *Kant's Copernican Revolution*. New York and Oxford: Oxford University Press.
- Brook, Andrew (1994) *Kant and the Mind*. Cambridge: Cambridge University Press.
- Castañeda, Hector-Neri (1999) 'I-Structures and the Reflexivity of Self-Consciousness'. In Hector-Neri Castañeda, *The Phenomeno-Logic of the I. Essays on Self-Consciousness* ed. James G. Hart and Tomis Kapitan (Bloomington, IN: Indiana University Press), 251–91.
- Dickerson, A. B. (2004) *Kant on Representation and Objectivity*. Cambridge: Cambridge University Press.
- Kant, Immanuel (1998) *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, [= CPR].
- (2002) *Theoretical Philosophy after 1781*, ed. Henry Allison and Peter Heath. Cambridge: Cambridge University Press.
- Miller, Alexander (2002) 'Introduction'. In Alexander Miller and Crispin Wright (eds), *Rule-Following and Meaning* (Montreal: McGill-Queen's University Press).
- Rosefeldt, Tobias (2007) 'Review of A. B. Dickerson, Kant on Representation and Objectivity'. *Philosophical Review*, 116/3, 468–70.
- Sellars, Wilfrid (1958) 'Counterfactuals, Dispositions, and the Causal Modalities'. In Herbert Feigl et al. (eds), *Minnesota Studies in the Philosophy of Science*, vol. 2 (Minneapolis: Minnesota University Press), 225–308.
- Thöle, Bernhard (1991) *Kant und das Problem der Gesetzmäßigkeit der Natur*. Berlin: de Gruyter.