4

The Unity of Kant’s Active Thinker

Patricia Kitcher

Introduction

Kant thought that he could establish conclusions about the mind and nature by making transcendental arguments; he believed that his results had anti-empiricist, if not anti-naturalistic, implications about the mind. The themes of transcendental philosophy, naturalism, and the mind are closely tied together in his work. This is no accident. It is largely because of our Kantian heritage that we think of the three issues as interconnected. Over the last forty years or so, many philosophers have believed that a Kantian approach could be employed in the service of systematic philosophy. I have three aims in the chapter. First, I try to clarify the Kantian legacy by considering more closely how these issues were connected in his work. Second, I draw on that analysis to show that Kant’s transcendental arguments had at least one anti-mechanistic and two anti-empiricist implications about the mind. Third, I argue that he achieved these results because he made stronger assumptions about the psychological prerequisites of cognition than most contemporary philosophers are willing to countenance. If I am correct, then some of Kant’s substantive theses may be as useful to contemporary epistemology and philosophy of mind as his distinctive transcendental method.

Section 1 gives a brief overview of how the transcendental deduction (hereafter ‘TD’) is supposed to work. The second and longest section presents an account of a central piece of the TD, the argument for transcendental apperception. I discuss this argument in some detail to
show that it involves features that are not shared by other transcendental arguments. In section 3 I bring out the strength and uniqueness of Kant’s argument for the unity of transcendental apperception by contrasting it with a very sophisticated contemporary attempt to show that object cognition requires a continuing thinker—that offered by Quassim Cassam. In the final section, I argue that in exploring the requirements of cognition, Kant discovers a kind of consciousness that both plays a crucial role in mental unity and is not obviously explicable in terms of current naturalistic models of mind. In David Chalmers’ (1996) terminology, it is a kind of consciousness that raises a ‘hard problem’ for contemporary theories, even though it is unrelated to standard cases of the ‘hard problem’, cases such as the unpleasant quality of pains or the ineffable visual quality of a purple haze. The last two sections thus argue that Kant’s TD has implications for contemporary work on the mind, specifically for theories of mental unity and for naturalistic theories of consciousness.

1. What is a transcendental deduction?\(^1\)

Transcendental arguments have interested recent philosophy because of their seeming potential to defang various sorts of scepticism. Two famous types of scepticism thought vulnerable to Kantian-type arguments were scepticism about external objects and scepticism about other minds. On the other hand, a number of historians of philosophy beginning (I believe) with Margaret Wilson (1974) have denied that Kant had Cartesian scepticism in his sights. His target was not Descartes, but Hume. Kant’s description of the problem for *The Critique of Pure Reason* supports the historians. His project was to explain

_How are synthetic a priori judgments possible?* (A19, cf. A9)\(^2\)

He defines *a priori* cognition in contrast to the *a posteriori* variety in the Introduction to the First Critique.

---

\(^1\) Some of the material in this section appeared in (Kitcher 2008).

\(^2\) References to the *Critique of Pure Reason* will be in the text, with the usual ‘A’ and ‘B’ indications of editions. In providing English translations, I usually rely on (Pluhar 1996), but I also borrow freely from (Kemp Smith 1968), and from (Guyer and Wood 1998), and sometimes I combine translational suggestions from the different standard references. Where I substantially alter a translation, I indicate that the translation is amended. I also use Guyer and Wood’s convention of indicating Kant’s emphasis through boldface type. References to Kant’s works, other than the first *Critique*, will be to (Kant 1900–) and will be cited in the text by giving volume and page numbers from that edition.
But even though all our cognition commences with experience, nevertheless, it does not for that reason all originate from experience. For it might well be that our empirical cognition itself is a composite of what we receive through impressions and of what our own cognitive faculties give up out of themselves (merely induced by sensory impressions). (B1)

One calls such cognitions [i.e., what our cognitive faculties give up out of themselves] a priori, and distinguishes them from the empirical [ones], which have their sources a posteriori, namely in experience. (B2)

Since the a priori elements of cognitions come not from objects, but from activities of the mind, there is a special problem in establishing the legitimacy of their use.

Kant believed that in order to vindicate the use of a priori concepts he needed to develop a new type of argument, a TD. He was explicit about the unique feature of such a deduction:

the transcendental deduction of all a priori concepts has a principle to which the entire investigation must be directed: viz., the principle that these concepts must be recognized as a priori conditions for the possibility of experience (whether the possibility of the intuition found in experience, or the possibility of the thought). (A94/B126)

In transcendental cognition, so long as we are concerned only with the concepts of the understanding, our guide is the possibility of experience... The [transcendental] proof proceeds by showing that experience itself, and therefore the object of experience, would be impossible without a connection of this kind [between concepts]. (A783/B811)

These descriptions immediately invite the question: What is the ‘possibility of experience’?

On this point, Kant’s published and unpublished remarks are clear. The possibility of ‘experience’ should be understood as the possibility of ‘empirical cognition’:

The categories serve only for the possibility of empirical cognition. Such cognition, however, is called experience. (B147, see also, e.g., 7.141, 18.318)

And by the ‘possibility of empirical cognition’, he means the possibility of attaining cognition of objects through receiving information through the senses.3

Kant signals that the assumption that cognition through the senses is possible for creatures like us is an acceptable starting place for his defence of

3 Carl (1989, 1992) argues against this view. See (Kitcher 2011: Ch. 7) for discussion.
the categories in the opening sentence of the Introduction (the sentence just preceding the material cited above):

There can be no doubt that all our cognition begins with experience [with the senses rousing our cognitive power to its operation]. (B1)

The assumption of the possibility of empirical cognition is a uniquely appropriate starting place for a defense of *a priori* cognition. He would be able to argue that *a posteriori* cognition, against which the *a priori* is unfavourably compared, itself requires *a priori* contributions from the mind.

We seem to have a simple account of the argument. It starts from the assumption of empirical cognition and then regresses to certain *a priori* concepts whose use is shown to be a necessary condition for the possibility of empirical cognition. This is too simple in some respects, but I will add only two complications. As Kuehn (1997) among others notes, at a crucial point in the set-up to the TD, Kant explained his project in terms of ‘objective validity’:

With the categories of the understanding we encounter a difficulty that we did not encounter in the realm of sensibility: viz. how subjective conditions of thought could have objective validity. (A89/B122)

Kuehn connects Kant’s description of his problem to the theories of Christian Wolff (1751/1983). Wolff understood philosophy as ‘the science of all possible objects, how and why they are possible’ (Kuehn, 1997: 229). On Kuehn’s account, Wolff’s goal was to prove that certain concepts were possible, by showing that objects falling under the concepts were possible (1997: 232). Kuehn notes that G. F. Meier’s logic text, which Kant used as a basis for his lectures, laid out the basics of the Wolffian programme:

A learned concept which has been created by arbitrary conjunction must be proved or disproved. We can achieve this either by (i) experience, if we show that their concepts are real or not real, or (ii) by reason, either directly or indirectly by showing that and how their objects can become real or that they cannot become real. (cited in Kuehn 1997: 235)

Suppose that Bart has the concept ‘ooblick’, the concept of green sticky stuff that falls from the sky like rain. He could prove the possibility of his concept either by seeing some ooblick (not likely) or by explaining how ooblick might come into existence. In light of the Wolffian background, we can understand why Kant had no worries about empirical concepts.
Their legitimacy or objective validity could always be established whenever a doubt arose (A84/B116–17). We can also see why sensitive readers such as Strawson thought that the success of the TD must imply not just the necessity of using object concepts, but the existence of ‘objects’ in the sense captured by the *a priori* concept of something that has properties and undergoes change through its interaction with other things.

The second complication comes from Kant’s famous analogy between the TD and legal deductions. Oddly, this clue remained unexplored until Henrich’s (1989) pioneering study. As he explains, the practice of deduction-writing arose because some means were needed to settle disputes about property and inheritance among the various entities that had constituted the Holy Roman Empire. He explained how they worked:

In order to determine whether an acquired right was real or only presumption, one must legally trace the possession somebody claims back to its origin. The process through which a possession or a usage is accounted for by explaining its origin, such that the rightfulness of the possession or the usage becomes apparent, define the deduction. (Henrich 1989: 35)

As he notes, the analogy with legal deductions explains the TD’s numerous references to the ‘origins’ of representations. In the preparatory section to the TD, Kant characterized one of its aims as providing for the categories ‘a *certificate of birth* quite other than descent from experience’ (A86/B199, my emphasis).

Kant assumed that there were only two possible origins for representations. Either they arose through outer causes operating through the senses or through inner causes, that is, through the activities of the mind (A98). The only alternative to an empirical birth certificate would be to trace the origin of the categories back to the operations of mental faculties. In particular, Kant tries to show that categorial concepts arise from activities of the mind that are necessary for any thinking at all. That is, on analogy with legal deductions, the TD traces the usage of certain concepts back to operations of the mind that are necessary in all thinking, thereby providing an origin for the concepts that is suitable to their role, which is that they are applicable to all objects of cognition. An empirical derivation could show that the concept in question applies to some objects, but a TD is supposed to show why the categories apply to any object that can be thought at all. This analysis of the argument agrees with Stroud’s (1968) seminal discussion. Stroud thought that Kant’s argument had a much
better chance of defending the necessity of using certain concepts than later transcendental arguments because of its generality. Its claim is not just that some concepts must be used if others are—but that the possibility of thinking itself implies the necessity of using such concepts.

2. The arguments for mental act-consciousness and mental unity in the TD

The TD proper, the second chapter of the Analytic of Concepts, does not mention all the categories, much less argue for them. Rather, it tries to establish two preliminary results that are necessary for the eventual argument for the categories in the Principles chapter: the unity of the thinker and the necessary agreement between concepts and intuitions. I’m going to focus on the first task, that of establishing the unity of ‘transcendental apperception’. In the Critique, Kant expressly distinguishes apperception from inner sense. We can get a better appreciation of his theory of apperception by briefly considering why he initially considered inner sense to be a key mental faculty. I begin with Locke’s well-known introduction of inner sense.

The other Fountain, from which Experience furnisheth the Understanding with Ideas, is the Perception of the Operations of our own Minds within us, as it is employ’d about the Ideas it has got; which Operations, when the soul comes to reflect on, and consider, do furnish the Understanding with another set of Ideas, which could not be had from things without: and such are Perception, Thinking, Doubting, Believing, Reasoning, Knowing, Willing, and all the different actings of our own Minds... (Locke 1690/1979: 105)

Although the notion of an internal sense was, I think, novel, one phenomenon that underlies its introduction was familiar. Many in the logical tradition maintained that people are aware of the cognitive acts they perform. Taking the most prominent example, the Port Royal Logic (Arnauld 1662) assumes that anyone can tell when he is judging, inferring, remembering, seeing, and so forth.

Since this phenomenon is much less remarked today, I will try to make it vivid with an example. Consider the premises of a simple inference:

All men are mortal.
Caius is a man.
Normal humans are aware both of the conclusion and of a movement of their minds from the premises to the conclusion. With no effort at all, they can distinguish acts of inferring from cases where they see no relation and have to be told what follows from the premises. This distinction would be obvious even if someone were to say the conclusion as quickly as the mind might infer it; it would be obvious even if a neurosurgeon could induce the production of sub-vocal speech, ‘Caius is mortal’, as quickly as the mind could infer that conclusion.

The phenomenon of mental act awareness justifies neither of Locke’s assumptions. It implies neither that being so aware supplies thinkers with a concept of ‘reasoning’ nor that the awareness is best modelled by analogy with the ‘outer’ senses. I highlight mental act awareness not to support Locke’s introduction of ‘inner sense’, but to draw attention to one phenomenon that it was meant to illuminate, a phenomenon that is central to Kant’s theory of thinking.

For much of his career Kant was an inner sense enthusiast. He lauds the power of inner sense in an early essay (The False Subtlety of the Four Syllogistic Figures, 1762) where he explains the difference between so-called ‘animal cognition’ and rational human cognition. He is criticizing one of his contemporaries (the logician Meier) who had claimed that animals use concepts. [Meier’s] argument runs like this: an ox’s representation of its stall includes the clear representation of its characteristic mark of having a door; therefore, the ox has a distinct concept of its stall. It is easy to prevent the confusion here. The distinctness of a concept does not consist in the fact that that which is a characteristic mark of the thing is clearly represented, but rather in the fact that it is recognized [erkannt] as a characteristic of the thing. The door is something which does, it is true, belong to the stall and can serve as a characteristic mark of it. But only the being who forms the judgment: this door belongs to this stable has a distinct concept of the building, and that is certainly beyond the powers of animals.

I would go still further and say: it is one thing to differentiate things from each other, and quite another thing to recognize [erkennen] the difference between them . . . (2.59–60)\(^4\)

Animals can differentiate things from one another—in the sense that they can behave differently with respect to them. But that does not imply that they have any understanding of how they differentiate the objects.

\(^4\) This translation is from (Walford and Meerbote 1992: 103–4).
The essay continues by offering a hypothesis about how humans are able to recognize characteristic marks as such and so have (distinct) concepts.

My current opinion is that this power or capacity is nothing other than the faculty of inner sense, that is to say, the faculty of making [zumachen] one’s own representations the objects of one’s thought. This faculty cannot be derived from any other faculty. It is, in the strict sense of the term, a fundamental faculty, which in my opinion, can only belong to rational beings. But it is upon this faculty that the entire higher faculty of cognition is based... (2.60, Walford and Meerbote 1992: 104)

Why couldn’t an animal just recognize the mark as such, why must it be able to think about its own representations? Kant’s view of rational cognition is that in applying concepts, rational animals know the basis or ground for the application—hence they must be aware of their own representations, because those are the grounds of the application. By contrast, animals differentiate things only ‘blindly’, without any idea of the basis of their differential behaviour. I stress this point, because the cognition that is the central topic of the First Critique is the rational cognition just described. Although it sounds somewhat oxymoronic, Kant’s quarry is thus rational empirical cognition.

Johann Nicolaus Tetens wondered how the representations of inner sense could be understood as representations in the same sense as representations of outer sense (Tetens 1777/1979: 1.7.45). He answers his own question as follows: as objects cause impressions on sensory organs that give rise to sensations that represent the objects, (mental) acts that result in changes in representations cause impressions on the organ, the mind or brain, and those impressions give rise to sensations—which represent the actions. In light of his analysis, Tetens accepts a criticism of the cogito that he attributes to Johan Bernard Merian (1732–1807): Descartes should not have said ‘I think’, but ‘I have thought’ (Tetens 1777/1979: 1.47). Below I suggest why Kant might have demoted inner sense—as Tetens explained it—to a ‘lower’ faculty.

Let us turn to the TD itself or at least to some pieces of it. In the A Deduction, Kant offers a complex account of the various acts of putting representations together—or synthesizing them—that are necessary for empirical cognition. The necessity of the unity of apperception is introduced in the course of his account of the third synthesis, that of recognition in a

---

5 I argue for this important point in Kant’s Thinker (2011: Ch. 9).
concept. Unlike most discussions, he uses an example to explore the requirements of concept application. The example is counting:

Without the consciousness that what we are thinking is the same as what we thought an instant before, all reproduction in the series of representations would be futile...If, in counting, I forget that the units that now float before my mind or senses were added together by me one after another, I should never know that a total is being produced through this successive addition of unit to unit, and so would remain ignorant of the number... (A103, amended translation)

As has often been noted, he believes that concept application involves the use of rules. In this case, the person must be aware of applying the counting rule to the units (perhaps a real or imagined stroke symbol): the first is designated by ‘1’, etc. But something further is required. To have rational cognition—to know the basis of his judgement ‘four’, for example—the counter must recognize that judgement as the result of his application of the counting rule to his sensory evidence. How does this happen?

In the further discussion of the example, Kant suggests that a cognizer must be conscious of his act of judging:

This number’s concept consists solely in the consciousness of this unity of synthesis. The very word ‘concept’ could on its own lead us to this observation. For this one consciousness is what unites in one representation what is manifold, intuited little by little, and then also reproduced. Often this consciousness may be only faint, so that we do not [notice it] in the act itself, i.e. do not connect it directly with the representation’s production, but [notice it] only in the act’s effect. Yet, despite these differences, a consciousness must always be encountered, even if it lacks striking clarity; without this consciousness, concepts, and along with them cognition of objects, are quite impossible. (A103–4, my emphasis)

This is a strong claim. Cognition of objects would be impossible without a consciousness of acts of combining. But it should be clear why he thinks that these acts must be conscious. If cognizers were not conscious of these acts, then they would not know the basis of their judgements, and so would fail to be (rational) cognizers. He will allow that thinkers do not have to pay much attention to individual steps, adding up the stroke symbols little by little in accord with the counting rule; still they must be conscious that the act of judging ‘four’ is based on carrying out these steps.

For Kant, rational cognition requires a kind of act-consciousness. In his published Anthropology lectures, he characterizes consciousness of mental acts in terms of ‘apperception’ and relates that faculty to the understanding:
he contrasts ‘apperception’ with ‘apprehension’, a consciousness of particular mental states through inner sense (7.134n.). We can get a firmer grasp on Kantian apperception if we consider why he would have rejected inner sense, as Tetens understood it, as the basis of rational cognition. According to Tetens, inner sense is a sense, because it records acts of thinking. But mere awareness that you have judged could not underpin the rationality of the judgement. For that, the cognizer must be conscious, not that he has judged or even that he is judging right now. He must be conscious of judging on the basis of evidence, of having applied the rule to the data. Failing this, rational cognition through concepts—judging—would be impossible.

Consider also the ‘Caius is mortal’ inference. On Tetens’ theory, a cognizer would know that he had inferred by being aware of the trace left by his act of inferring ‘Caius is mortal’. Even supposing that inferring and other mental acts have somewhat different ‘feels’ or ‘flavours’, so that a reasoner can tell that ‘Caius is mortal’ was an inference and not a perception or something learned through testimony, this would hardly be sufficient to make him a rational reasoner. To be capable of rational inference, the reasoner must be aware—as he makes the inference—of his act as being based on premises. The creation of impressions of mental actions in a Tetensian inner sense is too little and too late to contribute to the rationality of inference or judgement.

In the counting passage, Kant has the cognizer calling the acts of applying the rule to units ‘mine’. What is the justification for saying that the numbers were added together by me, for example? Kant later agrees with Hume that humans can see no constant self in the flux of mental states:

There is, in inner perception, a consciousness of oneself in terms of the determinations of one’s state. This consciousness of oneself is merely empirical and always mutable; it can give us no constant or enduring self in this flow of inner appearances. (A107)

He also disagrees with Locke that the mere consciousness of states can explain how different states belong to a single ‘I’. In the B edition of the

*Kant’s explanation concerns why different states should be understood as belonging to a common subject—what we might call the ‘togetherness’ of different states. He does not address (except inadequately through inner sense) why these states should be understood as belonging to a particular subject, namely me. That is, he does not adequately address what might be called the ‘mineness’ problem of mental states. See (Kitcher 2011: Chs. 1, 2, 9, and 15) for further discussion."
TD he is explicit about the problem with Locke’s theory (though he doesn’t mention Locke by name):

The empirical consciousness that accompanies different representation is intrinsically sporadic and without any reference to the subject’s identity. (B133)

Kant’s objection is that Lockean consciousness—the consciousness which is inseparable from thinking (Locke 1690/1979: 335)—is momentary or episodic. As such, it cannot provide a basis for representing a common subject.

So how can humans be aware of themselves as continuing, existing, thinkers? The key is their awareness of their acts of synthesizing:

Reference to the subject’s identity...comes about not through my merely accompanying each representation with consciousness, but through my adding one representation to another and being conscious of their synthesis. Hence only because I can combine a manifold of given representations in one consciousness, it is possible for me to represent the identity of the consciousness itself in these representations. (B133, my emphasis)

But how does being aware of these acts of synthesis enable the cognizer to represent his identity?

Kant takes the representation ‘I think’ to be a priori (B132). We have just seen the argument that it cannot be a posteriori: neither inner intuition nor a Lockean ‘accompanying’ consciousness (if those are different) provides any evidence of a continuing self. As with the a priori categorial concepts, the ‘I think’ is associated with an a priori principle, the principle that different representations must belong to a common self or thinker (A117, B132). No amount of empirical data could establish this principle. It is nonetheless possible for cognizers to recognize instances of it. To see how and also how consciousness of synthesis permits subjects to recognize their identity, we may return to the counting example.

The counter is aware of four stroke symbols to which he applies the counting rule, 1, 2, etc. When the understanding applies the counting rule to information contained in the sensory states that float before the mind, it recognizes that the antecedent of the rule is fulfilled, so the judgement ‘4’ can be made. But it also recognizes something else. Through being aware of its act of synthesis, it recognizes that it has made the judgement on the basis of applying the counting rule to representations contained in sensory states. Thus, it recognizes that the judgemental state could not exist without the sensory states. The judgemental state must belong with the
sensory states to a single consciousness. That is, the understanding recognizes that the mental states it combines and the combined state that results from the combination as instances of the ‘I think’ rule. Because a counter applies two rules, the counting rule and the rule of apperception, she does not merely form the representation ‘4’, she also represents the states and acts of which she is conscious as the states and acts of a single cognizer. Consciousness of the act of synthesis is crucial for rational cognition; without it, conceptual or rational cognition of objects is impossible. With it she is also able to recognize the unity of her consciousness. This is why Kant claims in the ‘anti-Locke’ passage that it is only through engaging in cognition that one can recognize the identity of one’s consciousness through time. He repeats this extraordinary claim in the B edition discussion of the Paralogisms:

We are acquainted with the unity of consciousness itself only by its being for us an indispensable requirement for the possibility of experience. (B420)

The B Deduction makes an even stronger claim about the relation between cognition and the unity of consciousness.

Synthetic unity of the manifold of intuitions, as given a priori, is thus the ground of the identity of apperception itself. (B134)

It is not just that, if the multifarious contents of sensory states could not be combined in judgements, then would-be cognizers would lack experience (empirical cognition). Nor is the claim even that, under these circumstances, they would not be able to think about themselves as continuing cognizers. It is rather that, absent the combinability of intuitions, the identity of apperception would also be absent. As he explains slightly later:

All representations given to me must stand under this [original synthetic unity of apperception], however they must be brought under it through a synthesis. (B136)

That is, the unity of self-consciousness is brought about through combination. In engaging in cognition, the understanding also partially creates a rational subject.\(^7\)

\(^7\) See also (A112): ‘Without such unity [as produced by the rule for cause and effect, for example] no thoroughgoing and universal and hence necessary unity of consciousness would be encountered in the manifold of perceptions.’
I hope that I have laid out enough of the argument of TD that its structure and principal theses are clear. The argument regresses from the possibility of rational empirical cognition to the unity of apperception; having shown how thinking must work for creatures like us who acquire knowledge through combining information gained through their senses, Kant is then in a position to consider what the possibility of thinking itself requires. And what it requires is that the materials received through the senses can be combined in ways that make thinking and so the unity of apperception possible. He claims, but does not argue, that those ways of combining materials are captured in the *a priori* categories. Even at this point, however, we can see why he makes the strong and unusual claim that object cognition and the unity of apperception are necessary and sufficient conditions for each other (for example, B137). Failing the combination of some representations in a further representation such as a judgement or the conclusion of an inference, a cognizer would be unable to recognize the relation of necessary connection across her states that stands behind the use of ‘I think’; failing the ability to apply the ‘I think’ across her representations, a cognizer would lack rational empirical cognition, because she would not see some of her representations as the grounds of her judgements.

**3. Arguing for mental unity**

We can see how strong an argument Kant is able to mount against those who deny the legitimacy of an *a priori* concept of ‘person’ or ‘unified cognitive subject’ by contrasting it with one of the best recent attempts at presenting a ‘transcendental argument’ for the same conclusion, that of Quassim Cassam (1997). The essentials of Cassam’s wide-ranging account can be captured as follows. Neo-Kantians have argued that it is impossible to think of one’s experience as containing objects in the weighty sense (items that can be perceived and that can exist unperceived) unless one can self-ascribe perceptions and grasp the identity of the thing to which these perceptions are ascribed (Cassam 1997: 36). Because of the resemblance of this line of reasoning to Kant’s argument from object cognition to the unity of consciousness, Cassam refers to it as the ‘objectivity requires unity’ (ORU) argument. He notes that those who argue that there are no continuing persons will not be impressed. Theirs is a thesis about what
those objects who were thought to be persons really are; ORU concerns
the ways in which individuals must think of themselves in having cogni-
tion (Cassam 1997: 178–9, 181). He presses on, because he thinks that,
thanks to ORU, there is something to be explained—namely, how
cognizers use the term ‘I’—that may require reference to persons.

At this point, Cassam appeals to Gareth Evans’ (1982) analysis of the
range of capacities that are required for individuals to have ‘I thoughts’:
such individuals must be able to recognize the connection between
‘I thoughts’ and their special ways of gaining knowledge of their mental
states and physical properties, must recognize the connection between
‘I thoughts’ and behaviour, and so on (Cassam 1997: 189). So the argu-
ment is that objectivity requires unity, including the possession of
‘I thoughts’ and—moving from epistemological considerations to the
grounding of cognitive capacities in a subject—‘I thoughts’ can be had
only by creatures with various further capacities, who are thus persons, or
substantial subjects among other items in the (physical) world (Cassam 1997:
196–7). As Cassam concedes, this argument is weaker than it might be.
It avoids the fallacy of arguing from what cognizers must think they are to
what they are, by making a large assumption: only substantial subjects can
have the capacities required to be thinkers of ‘I thoughts’ (Cassam 1997: 197).

Kant’s version of the argument from (rational empirical) cognition to
unity of consciousness includes an element that is missing from the neo-
Kantian repertoire. To return, again, to the counting example, a counter is
aware of four stroke symbols to which he applies the counting rule, 1, 2,
etc. When he applies the rule, he recognizes that the antecedent of the rule
is fulfilled, and so judges ‘4’. Because he is at least implicitly aware of the
act of synthesis, he recognizes that he has made the judgement on the basis
of applying the counting rule to representations contained in sensory
states—and that his act of judging thereby creates a relation of necessary
connection across the sensory states and the judgemental state. Through
the self-conscious act of judging, the judgemental state comes to stand
in the relation of rational dependence to the sensory representations;
from the other direction, through that act, mental states of ticking off
‘1’, ‘2’, etc., achieve the status of grounds of cognition through the rational
dependence of the judgemental state upon them. Consciousness of the act
of synthesis is crucial for rational cognition. Without it, conceptual or
rational cognition of objects is impossible, because cognizers would not
know the bases of their cognitions. With that consciousness, however, the
cognizer creates a relation of rational dependence across his states in part by being at least implicitly cognizant of that relation. Since cognizers come with an *a priori* representation ‘I think’ that they apply according to the rule of necessarily belonging together, they can always attach ‘I think’ to the representations that participate in rational cognition.

Perhaps the view I am attributing to Kant will be clearer if we return to Stroud’s classic criticism of contemporary transcendental arguments. In Stroud’s view (1994: 234), Kant tried to legitimate the categories by arguing that they were both indispensably necessary for empirical cognition and objectively valid—they were true of objects of experience. The fundamental error of the neo-Kantians was to assume that the first project sufficed for the second. It could not, because the first project concerned only what cognizers must believe (and so what concepts they must possess) and there is no legitimate way to argue from the necessity of a belief in Xs to the objectivity validity of the concept ‘X’. Stroud’s criticism is widely recognized as sound and has shaped many subsequent discussions, including Cassam’s. He was also right about Kant’s intentions. In the case of the representation ‘I think’, the aim was not just to show that this *a priori* representation was necessary for rational empirical cognition, but that rational cognition required the existence of thinkers whose states were related as specified by that representation. And that is what Kant’s theory implies. Rational cognition does not just require that cognizers have an *a priori* representation ‘I think’. In any case of rational cognition, a cognizer must grasp, in making the judgement, that the relations of epistemic dependence specified in the rule attached to the ‘I think’ are fulfilled. A judgement is rational only when it has that genesis. Hence the *a priori* representation ‘I think’ and its associated rule play an essential role in producing rational judgements and thus rational judges or thinkers.

In this respect, the representation ‘I think’ is unique. Having a concept of external objects or of other minds is neither a necessary nor a sufficient condition for the existence of external objects or of other minds. But—when given suitable materials to work on and mental act awareness—the concept of a thinker whose representations stand in relations of necessary connection is a necessary and sufficient condition for a cognizer to form a rational judgement and so to be a thinker. In this case alone, the

---

* See also (Stroud 1968).
* Cassam (1987) discusses Stroud’s argument at some length. See also his (2003).
representation (plus other representations and the mental activity of combining them) is able to create the reality.

To move beyond the world of belief and representation to that of reality, Cassam appeals to the fact that various mental capacities require the use of a body. The representation ‘I think’ required by cognition of objects can therefore be deployed only by embodied or substantial subjects. Oddly, Kant is able to move from representation to reality by staying within the realm of cognizing and thinking.

4. Consciousness of mental acts

One important part of Kant’s account of rational cognition is the claim that cognizers must be conscious of their mental acts of judging and inferring. What kind of consciousness is this? As noted, Chalmers has distinguished easy kinds of consciousness, those that can be modelled, for example, by a computer, from the hard kinds, those that seem to resist modelling by any mechanism that is currently available. For many years the view has been that sentience, the ability to feel and sense, raises the hard problem. By contrast, sapience, the ability to think and represent, was easy. After all, the first computer models of mental processes were devoted to proving logical theorems and solving word problems. Apparently, these tasks could be carried out by simple production systems, that is, sets of conditional rules that tell the machine what formula to write down depending on the program and the current state. On Kant’s view, however, a simple production system could not model rational cognition.

Kant maintains that such cognition requires mental act awareness. It is tempting to think that a familiar complication to simple production systems could handle the needed awareness. Besides carrying out various rewrite rules, the program would contain a function that monitored its

10 Recently Christopher Peacocke (2008: Ch. 7) has started to investigate mental act-consciousness. His approach is to model mental act-consciousness on physical act-consciousness. Prima facie the cases seem disanalogous, because mental act-consciousness is essential to the performance of certain kinds of mental acts, such as judging and inferring. This doesn’t seem to be the case with physical actions, unless a physical motion can be understood as ‘action’ only when embedded in a rich description of mental activity. In that case, however, what needs to be understood is the mental activity.

11 Some, notably John Searle (1980), have argued that sapience is also a hard problem, but the ‘standard’ view is that the hard problem arises mainly in the case of sentience.
own rewriting activities. To evaluate this suggestion, I turn to Ned Block’s (1995) account of models of monitoring consciousness. Block offered three ways of modelling monitoring consciousness: (1) as metacognition, where being in a mental state is accompanied by the thought that one is in that state; (2) as self-scanning; and (3) as a phenomenal consciousness, where one is aware of some sort of phenomenal quality. Perhaps in this case it would be the feeling of judging.

None of these models seems to fit Kant’s description of judging or concept application. We have already seen the difficulty with (1). Rational cognition is not simply a matter of knowing that you are in a state, say, of judging, but of being conscious of entering that state through a synthetic act. The second option seems equally unpromising. If a sequence of mental states did not involve an awareness of entering the latter state through a synthetic act, then scanning that sequence cannot make up for the lack of this awareness. The third option seems even less plausible. Even if we assume that different mental actions have different feels, judging and inferring do not seem to be centred on phenomenal consciousness in the way that the paradigm cases of feeling pain or seeing purple are. If there is a phenomenal feeling to these acts, it is not their essence, which concerns the ability to recognize the relations across one’s own thoughts.

One symptom that a type of consciousness falls in the ‘hard’ category is the susceptibility of putative accounts to ‘Zombie’ objections. Thinking about Zombies (those who act just like people, but lack consciousness) may make it clearer why Kant downgraded ‘inner sense’ (on Tetens’ theory of what it involves). Insofar as Tetens is right about how inner sense works, then inner sense is not sufficient for explaining rational cognition. Creatures with just inner sense would be cognitive Zombies. They would look and sound quite like cognizers, but they would not be rational cognizers. Such creatures would process information or, in Kant’s terminology, combine representations in regular ways. They might also have metacognitive beliefs about the information (or representations) that they have received. They might be aware that some of their representations have the feel of judgements, so they could report which of their beliefs come from their own processing or combining as opposed to being received from others. Finally, they might even be aware of which of their representations are connected to their judgements as their bases. In Kant’s view, even if such Tetensian creatures could say which of their representations were judgements and could connect these to their bases, they would
still not be rational cognizers, because they could not make the connection in the right way. They could not see their judgements as the results of acts of judging on the basis of the perceptual evidence. Alternatively, although they would know in a sense that these are judgements that they made rather than received, they would not see any fundamental difference between thinking for themselves and taking on the opinions of others. Still a third formulation: although they would connect judgements to their bases, the connections would seem accidental, because they would not understand how the judgemental representations were dependent on the earlier representations. The problem with modelling Kant’s mental act awareness as monitoring consciousness is that the latter is a contemporary reincarnation of Locke’s reflective consciousness, the consciousness that does or can accompany mental states. Kant saw clearly that an episodic reflecting consciousness could not account for mental unity, the relations of necessary connection across mental states. Since reflecting or monitoring consciousness could not account for the relations of epistemic dependence across the states of a rational cognizer, it could not explain rational cognition either. These intertwined inadequacies of monitoring consciousness are just what led Kant to introduce a new kind of awareness, mental act-awareness or apperceptive consciousness.

References


Kant, I. 1900–. *Kants gesammelte Schriften, Akademie Ausgabe*. Edited by the Koeniglichen Preussischen Akademie der Wissenschaften, 29 vols. Walter de Gruyter and predecessors.


