

ARTICLE

## Kant's "I think" and the agential approach to self-knowledge

Houston Smit

Department of Philosophy, University of Arizona, Tucson, USA

### ABSTRACT

This paper relates Kant's account of pure apperception to the agential approach to self-knowledge. It argues that his famous claim 'The **I think** must be able to accompany all of my representations' (B131) does not concern the possibility of self-ascribing beliefs. Kant does advance this claim in the service of identifying an a priori warrant we have as psychological persons, that is, subjects of acts of thinking that are imputable to us. But this warrant is not one to self-knowledge that we have as critical reasoners. It is, rather, an a priori warrant we have, as thinkers, to prescribe to given representations their conformity to principles of thinking inherent in our capacity of understanding itself.

**KEYWORDS** Self-knowledge; self-consciousness; apperception; understanding; person; transcendental deduction of the categories; principles of thinking

Directly and indirectly, Kant's critical philosophy has inspired an important approach in contemporary theorizing about the nature of our self-knowledge, one to which I will refer as 'the agential approach'. The self-knowledge in question is a distinctive sort of knowledge that each of us has, *de se*, of her own beliefs. On the agential approach, one has this self-knowledge in virtue of being a critical reasoner. To be a critical reasoner, in turn, is to have the ability to adopt beliefs out of one's own recognition of how reasons dictate that one ought to. And a critical reasoner has this ability only if her conceptual awareness of herself as exercising this control and agency over her thinking is correct and justified. Proponents of the agential approach contend that a critical reasoner has the right, or a priori warrant, to take herself to in fact have the thoughts and attitudes that do, or can, enter into her critical reasoning, because, if she did not have this right, she would not be subject to rational norms in the way that is constitutive of her being a critical reasoner. It follows that each critical reasoner has, as a critical reasoner, a certain kind of nonobservational knowledge of her own beliefs.<sup>1</sup> In the case of at least some proponents of the agential approach, Kant's account of self-consciousness,

**CONTACT** Houston Smit  [smit@email.arizona.edu](mailto:smit@email.arizona.edu)  Department of Philosophy, University of Arizona, Tucson, USA

© 2019 Canadian Journal of Philosophy

and in particular of pure apperception and the 'I think', seems to have played a role in inspiring them to take this approach.<sup>2</sup>

In what follows, I examine whether, and if so just how and to what extent, the account of self-consciousness Kant advances in the first *Critique* is continuous with the agential approach to self-knowledge. And in doing so, I will be focusing on Kant's famous claim 'The **I think** must be able to accompany all of my representations' (B131), one to which many proponents of the agential approach allude as expressing the core idea of behind this approach.<sup>3</sup> My goal in doing so is not to question the Kantian heritage of the agential approach. It is, rather, to further our understanding of Kant. For it will prove illuminating to see that Kant does indeed develop his account of self-consciousness – in particular, of pure apperception and its original synthetic unity – in the service of articulating and establishing an a priori warrant that is, in a crucial respect, of the same sort as that with which proponents of the agential approach to self-knowledge are concerned: namely, an a priori warrant that each of us has simply as a subject that is able to engage in thinking that is responsive to grounds of thought and cognition in the way that is characteristic of persons. This is a deep point of continuity between Kant, on the one hand, and proponents of the agential approach, on the other.

At the same time, two fundamental and closely related points of discontinuity will emerge. First, the a priori warrant that Kant aims to establish is not, I will argue, one specifically for claims to self-knowledge, or even necessarily for knowledge claims at all. It is, rather, a *de se* a priori warrant that each of us has, as a person, to put her capacity of understanding to use in acts of thinking that are imputable to her.<sup>4</sup> Now, on Kant's account, our capacity of understanding is conceptual, where a concept is a rule that constitutes the predicate of a possible judgment. And we have a genuine capacity of understanding at all only if we can correctly and legitimately take manifolds of representations that are given to us as ones that are subject to concepts. Indeed, Kant maintains that our capacity of understanding itself is the capacity to prescribe, correctly and legitimately, to particular manifolds as they are given to our consciousness – where the manifold may be one of concepts or the manifold of a given sensible intuition – their objectively necessary conformity to certain fundamental purely intellectual principles of our thinking. The categories number among these principles. Second, the a priori warrant that Kant sets out to establish does not, as on the agential approach to self-knowledge, have its ultimate ground in the nature of our cognitive agency in making inferences that are subject to rational norms. Rather, as we will see, Kant argues that the a priori warrant I have for any and all uses of my capacity of understanding has its ultimate ground in the nature of the cognitive activity that realizes in me my capacity of pure understanding. This activity is the operation of the spontaneity of cognition that first brings manifolds of representations given to me in intuition under certain purely intellectual

principles of our thinking, the highest of which is the principle of the synthetic unity of apperception.

### Transparency and the agential approach to self-knowledge

Many philosophers have espoused the idea that a subject's beliefs are transparent to her: that the question whether one believes a certain proposition *P* is to be answered, not by introspection, but simply by thinking about whether *P* is true.<sup>5</sup> Different philosophers have been led to this idea in rather different ways. But a number of those who champion the doctrine of the transparency of belief take their inspiration, at least in part, from Kant. Gareth Evans, in particular, espouses this doctrine in developing his influential account of the self-ascription of belief (Evans 1982, 225–6). In doing so, he makes reference to what he takes to be Kant's views about the 'I think'. According to Evans, a subject can use 'I believe' or 'I think' in attributing a belief to herself only if she can, in conjunction with physical predicates, use cognates of 'believe' and 'think' to attribute beliefs to others (Evans 1982, 226–8).<sup>6</sup>

Richard Moran retains the core idea of Evans' account of the transparency of beliefs in his influential development of the agential approach to self-knowledge (Moran 2001). On his view, beliefs are essentially states that a critical reasoner can author in herself on the basis of reasons that this subject can determine, in an exercise of her rational agency, to be reasons that tell in favor of her adopting that belief. Moran maintains that, to the extent that the beliefs we are to ascribe to a critical reasoner are, in this way, ones that are to be imputed to her, they must be transparent to her. For a critical reasoner, as such, has the right to take her beliefs to be subject to her critical reasoning, and thus to take reasons she has to hold a belief as evidence that she in fact holds that belief. This right is the basis of the distinctive sort of self-knowledge that, on Moran's view, any critical reasoner as such can, and must be able to, have of the beliefs that belong to her, merely by thinking about the reasons she has for adopting those beliefs.

But the agential approach to self-knowledge can be pursued without retaining the doctrine of the transparency of belief. Consider here Tyler Burge, to whom we owe the richest account on offer of the nature of the epistemic right to self-knowledge that, on the agential approach, each of us has as a critical reasoner. In Burge's terminology, 'entitlements are epistemic rights or warrants that need not be understood by, or even accessible to, the subject' (1993, p. 458); an individual's epistemic entitlement, unlike her epistemic justification, 'consists in a status of operating in an appropriate way in accord with norms of reason, even when the norms in question cannot be articulated by the individual with that status' (Burge 1996, 93). On Burge's view, each of us, as a critical reasoner, has a certain sort of a priori entitlement to claims, made in a distinctively first-personal way, about her

own beliefs, which entitlement is grounded, at least in part, in what is necessary to our being critical reasoners: one is a critical reasoner only if one's thinking is subject to rational norms that govern how one is to change one's beliefs based on one's critical reflection, but one is subject to rational norms in this way only if one has a certain a priori entitlement to take one's judgments about one's own thoughts to be true, which entitlement is 'stronger than that involved in perceptual judgments' (Burge 1996, 98).

Burge, however, denies that a critical reasoner's beliefs, even those that are hers because she arrived at them through her reasoning, are transparent to that reasoner. He does so on the grounds that a critical reasoner's ability to review her own reasoning critically, an ability essential to her being a critical reasoner, also requires that she be able to distinguish between how she ought to reason and how she, in fact, has reasoned. On Burge's account, then, a critical reasoner must have another mode of epistemic access to how she in fact reasons, in addition to the distinctively agential and nonobservational one she has in and through thinking about how she ought to reason.<sup>7</sup>

### Is Kant a precursor to the agential approach to self-knowledge?

Now one might well be tempted to take the agential approach to self-knowledge, and indeed even the transparency thesis, as coming to expression in Kant's famous claim, mentioned earlier, about 'the **I think**'. This claim (hereafter, 'the Claim') opens Section 16 of the Transcendental Deduction of the Categories in the B-Edition of the *Critique of Pure Reason*:

The **I think** must **be able** to accompany all my representations [*muss alle meine Vorstellungen begleiten können*]; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible, or at least nothing for me. (B131-2)

Here Kant asserts that the possibility of the <I think> accompanying a representation is a necessary condition of that representation being mine. A representation's being mine, in the relevant sense, is incompatible with my not being able to think what that representation represents, an inability that entails, in turn, that representation's being, at best, 'nothing for me'. It isn't obvious just what Kant is saying. But it seems plausible that a representation's being mine requires that I be able to put it to some use in my thinking, and in particular in thinking in which I am responsive to reasons in the right way. And it is certainly tempting to interpret what it is for the <I think> to actually accompany a representation in me to consist in my self-ascribing that representation. What it is more, one could make the case that what Kant here requires, as a condition of a representation's being mine, is the possibility of a self-ascription that constitutes knowledge of one's own representations. For a representation to have a use in my thinking, I must, in fact, have that representation. And it

seems plausible that Kant advances the Claim in the service of identifying an a priori epistemic right or warrant that each of us has, merely as a thinker, to ascribe to herself representation that she puts to use in her thinking to herself as ones that do, in fact, have a use for her in her thinking. On such a line of interpretation, then, the Claim advances a position strikingly similar to that taken by proponents of the agential approach to self-knowledge. It parallels the contention, made both by Burge and Moran, that a critical reasoner must be able knowledgeably to self-ascribe any beliefs she has as a critical reasoner.<sup>8</sup>

To be sure, even if we adopt this reading of the Claim, it isn't clear just how strong its continuity with the agential approach to self-knowledge is. In particular, is the conception that Kant is invoking of what it is for me to be a thinker, and thus a subject that can have a representation in the way I take myself to in calling that representation mine, the same as that which informs the agential approach to self-knowledge? It is tempting to answer this question in the affirmative. There is, however, an obvious point of discontinuity between the Claim and the position advanced by proponents of the agential approach to self-knowledge that should immediately give us pause. The Claim asserts the necessity of the possibility of the *<I think>* accompanying *all my representations*. And Kant famously provides a taxonomy of all our representations that divides all our representations with consciousness, or *perceptiones*, into two fundamental kinds, cognition and sensation, and then cognition, in turn, into concept and intuition (A320/B376).<sup>9</sup> The Claim, then, would seem, on the face of it, to hold for sensations and intuitions, as well as for thoughts and attitudes. Indeed, once we have examined Kant's conception of what it is for the *<I think>* to accompany one of my representations, we will be in a position to see that even *my* sensations, as *my* representations, must, on Kant's account, meet conditions of the possibility of being accompanied by the *<I think>*. Burge and Moran, by contrast, explicitly restrict their account of self-knowledge to one's knowledge of one's own thoughts and attitudes, to the exclusion of our sensations.<sup>10</sup> This raises the suspicion that the concept of what it is to be a thinker that motivates Kant's claim that the *<I think>* must be able to accompany all my representations cannot be entirely the same as that which informs the agential account of self-knowledge.<sup>11</sup>

### **Kitcher's answer**

In her recent work, Patricia Kitcher answers our last question in the affirmative. She develops an original reading of Kant's account of the transcendental unity of apperception on which it amounts to an account of rational thinking of the same sort that underlies the agential approach to self-knowledge. On her reading, the transcendental unity of apperception is a unity of a subject's mental states that constitutes him as a subject of what she dubs 'rational cognition'. In her parlance, rational cognition is

cognition of the sort humans have in and through making judgments and drawing inferences – where this is understood, in turn, as performing acts of ‘making and recognizing relations of dependence and so necessary connection across their states’ (Kitcher 2017, 170). As Kitcher reads him, Kant maintains that a thinker, in consciously drawing inferences from her various mental states, thereby creates and recognizes ‘the relation of necessary connection across them that makes them the states of a single thinker’ (Kitcher 2017, 171). The necessary unity of apperception, on Kitcher’s interpretation, is the necessary unity that a thinker realizes in her mental states when she uses them, in her rational thinking, to produce her beliefs. Kitcher grants that Kant himself held that a thinker self-ascribes representations only in being aware of them through inner sense, and so through a sort of observation of her own mental states. Nonetheless, she contends, his account of higher cognition is consonant with the agential approach to self-knowledge. Since judging and inference require conscious acts, ‘whenever a subject is consciously thinking that *p*, she can self-ascribe the activity and the thought’ (Kitcher 2017, 171): the consciousness had in consciously believing *p*, gives her grounds for self-ascribing this belief. But, she claims, Kant’s doctrine of the unity of apperception is, at the same time, incompatible with Evans’s transparency thesis, because it makes all of a subject’s beliefs, including those about the world, dependent on a prior *de se* consciousness she has of herself, in engaging in rational thought, as the subject that engages in this thought.<sup>12</sup> Finally, she brings this reading to bear specifically on the Claim by suggesting that the scope of the Claim, and indeed more generally of what she terms Kant’s ‘I-think doctrine’, is restricted to ‘the set of representations that can participate in cognition’ (Evans 1982, 144). On her reading, what makes an intuition that is in me a rational cognition, and so mine, is my being able to bring it under a concept so as to ‘combine it with other representations in a resultant representation’ (ibid). Kitcher contends that, in this way, Kant’s position is opposed to Evans’s, in that Kant’s position on higher cognition is incompatible with the doctrine of the transparency of belief: contrary to what Evans thought, Kant was ‘no friend of transparency, but its natural enemy.’<sup>13</sup>

Kitcher’s reading has many virtues, not least that it represents a serious, and philosophically resourceful, attempt to identify a single cogent, and well-motivated, account of self-consciousness and cognition in the many bewilderingly difficult texts in which Kant presents this complex account. Moreover, as advertised, Kitcher’s reading recognizes, and tries to do justice to, how Kant’s account of the possibility of our thought and cognition makes the synthetic unity of apperception explanatorily fundamental (cf. Kitcher 2011, especially Chapter 9). And in doing so, it relates Kant’s work to contemporary debates, including those about self-knowledge, in intriguing ways. Nonetheless, we have reason to worry that Kitcher’s reading fundamentally mischaracterizes

Kant's account of pure apperception and its synthetic unity. Indeed, once we have examined this account, we will see that we have reason to question an assumption made, not just by Kitcher (as well as other proponents of the agential approach to self-knowledge who look to Kant as an ally), but by most of Kant's readers: namely, that in accompanying one of her representations with the <I think>, a thinker is, whatever else, doing something above and beyond what she is doing in using that representation to think of what that representation is a representation of; and that one can, thus, put a representation to use in thinking without actually accompanying it with the <I think>. Examining the Claim, and the one-clause argument for it that immediately follows it, will suggest otherwise: any act of thinking in which a thinker relates one of her representations to an object consists, in respect of its intellectual form, in the <I think> accompanying that representation. But before we are in a position to examine the Claim, I need to take care of some preliminaries. The first is to explain, if only briefly, some of the terminology that it employs, as well as the treatment of the concept of combination in Section 15 that sets the stage for Section 16.

### Representation, thought, and combination

Kant characterizes representation (*Vorstellung*) as 'what has a relation to an object [*eine Beziehung auf ein Object hat!*]' (cf., e.g. 24: 805). I take this to amount to the characterization of representation as what the subject of a capacity to represent (which capacity may be that of a brute animal) *is to relate* to an object in exercising that capacity. Here – and unless I specify otherwise, in what follows – I will use 'object' to translate Kant's Latinized '*Object*'. '*Gegenstand*' is another term in Kant that is often translated with 'object'; but a *Gegenstand*, in his sense, is the real – the subject of activity and power – considered insofar as it is given as such in representation. In the case of any conceptual understanding, Kant holds that a *Gegenstand* can be given to it only in an operation of its ability of sensibility. An object, in his sense, in turn, is whatever a subject is conscious of, insofar as that subject is conscious of it (A189/B234). Since our capacity of understanding is conceptual, a *Gegenstand* that is given to you constitutes an *Object*, only insofar as you are conscious of it.

Any representation can be considered as an operation of a subject's capacity or ability of representation; to do so, in Descartes' famous terminology, is to consider it in its formal reality. But what makes it a representation, on Kant's account, is its being such that this subject is to relate it to an object, and to consider it insofar as it has this relation is to consider the representation in its objective reality. For example, Kant characterizes a sensation 'as an effect of an object [*Gegenstand*] on the ability for representation insofar as we are affected by it' (A19/B34). This is a sensation's formal reality. But what makes such an effect a representation is, in the case of an objective sensation, its providing the matter

of appearance: as such matter, an objective sensation (the green of the field) is to be related to an object (*Gegenstand*) of experience (the field) in an act of thinking that does not constitute an act of cognizing that object (*Gegenstand*).

Thinking (*Denken*), in Kant's sense of the term, is representing through concepts. And for a representation to be mine – that is, to belong to me *qua* thinker and indeed as the subject to whom acts of thinking are imputable and so as what Kant calls a psychological person – is for it to be one that I *am* to relate to an object, somehow, in an act of thinking. So, for example, in making the perceptual judgment 'This rose is red', I relate to an object, in this act of thinking, not just all the representations that make up the concepts <rose> and <red> as they make up these concepts (i.e. as general marks, marks that are common to more than one possible thing), but also those that make up my present empirical intuition (e.g. the perceptually presented token instances of shapes characteristic of a rose, which, as singular marks, are not common to more than one thing). And, in relating these representations to the rose in this judgment, I think, not just these concepts, but also this empirical intuition, along with the objective sensation <red> that belongs to this intuition. To be sure, the sensation <red> does not, of itself, have any relation to the rose; it is only insofar as it constitutes the matter of an appearance of the rose that the sensation has any relation to the rose. And I think it as having this relation when, in making this perceptual judgment, I subsume it under the concept <red>, the concept of the quality (redness) of the sensation that, as a 'predicate of appearance,' 'can be attributed [*beigelegt werden*] to the object [*Object*] in itself, in relation to our sense' (B69-70n).

Kant maintains that to think 'is to unite representations in a consciousness,' and that 'the unification of representations in a consciousness is judgment' (4: 304). Moreover, he distinguishes different actions of combination according to the unity of representations in a consciousness that it determines as its effect. He terms the most general such action 'combination in general [*Verbindung überhaupt*],' and examines the possibility of this combination in Section 15 of the B-edition Deduction (entitled 'On the Possibility of a Combination in General'). Kant opens this section with the claim that 'Only the combination [*die Verbindung*] (*conjunctio*) of a manifold in general can never come to us through sense ...; for it is an *actus* [*Aktus*]<sup>14</sup> of spontaneity of the power of representation' (B130). Kant tells us that he gives this 'act of the understanding [*Verstandeshandlung*]' 'the general title **synthesis**' to indicate

that we can represent nothing as combined in the object [*im Object*] without having previously combined it ourselves, and that among all representations **combination** is the only one that is not given through objects [*Objecte*], but can be executed only by the subject itself, because it is an *actus* [*Actus*] of its self-activity. One can here easily see that this action must originally be unitary [*einig*] and equally valid for all combination ... (B130)



Any token act of thinking is, on Kant's account, an act of combining (uniting representations in a consciousness) in which the subject is conscious of a manifold of representations given to her through an object, and in doing so represents that manifold 'as combined in' that object. *My* determined thinking is thinking in which I determine my representation of how the manifold is combined in the object out of my consciousness of how a principle of thinking requires me to do so. Here, and throughout what follows, I use 'determine' in Kant's technical sense, on which to determine is 'to posit a predicate to the exclusion of its opposite' (AA 1: 139). On Kant's account, any act of determined thinking consists in the 'originally unitary' action of 'combination in general' insofar as the subject combines in this action the manifold that an object could through sense supply for this action, which manifold constitutes the content that the subject of this action is subsequently to relate to that object in *her* determined thinking. 'Combination in general,' out of which any act of determined thinking consists, is thus the subject's self activity in thinking considered in itself and so insofar as it is not, itself, determined.

I am proposing, then, that not every act of determined thinking that occurs in me is, on Kant's account, an act of what Kant calls '*my* determined thinking' (B134). As we will see, *my* determined thinking is that in which I relate some particular manifold of representations that is given to my consciousness to an object in and through prescribing to that manifold its conformity either to a logical form of judgment (in the case of a manifold of concepts) or to a category (in the case of the manifold of an intuition). For it is only in such prescription that I posit a unity of the given manifold in the object to the exclusion of its opposite, and this by correctly and legitimately representing some ground (not necessarily in the object) as one sufficient to determine this unity. *My* determined thinking constitutively employs the most fundamental concepts of an object in general [*Object überhaupt*] that are inherent in the discursive capacity of understanding of the kind we have, as purely intellectual principles of our thinking. These concepts, Kant argues, are the twelve categories that he exhibits in the Table of Categories.

Our main aim in the remainder of this paper will be to present, if only schematically, how Kant argues, over the course of Sections 15 and 16, that *my* determined thinking is made possible only by pure apperception. Pure apperception is the operation of understanding that combines all manifold of intuition that is to be encountered in me in the concept of an object in general. Indeed – and this is the crucial point – pure apperception combines this manifold as the act that first subjects these manifolds to the categories, an act that Kant terms 'the legislation for nature' (A127). Pure apperception is thereby the operation of my capacity of apperception that realizes it as my capacity of understanding, my first capacity for determined thinking.

Moreover, in legislating for nature, pure apperception also subjects all other manifolds of representations that are to have any use in my thinking to principles of thinking. Pure apperception is thus the act that realizes my capacity of apperception more generally as the power to unite any manifold given to me in combination in general in an a priori consciousness of how this manifold is subject to purely intellectual principles of my thinking. And, in doing so, pure apperception is what constitutes me as the numerically identical subject throughout any manifold of representations given to my consciousness that are to have a use in my determined thinking, and so that are to be anything to me (B134). As such pure apperception is not, and cannot be, itself an instance of *my* determined thinking. It is the single, and original, purely intellectual de se consciousness that first makes any of my determined thinking possible by subjecting a manifold given to my consciousness to intellectual conditions to which it must conform, if the representations that make up that manifold are to have any use in my determined thinking. And, on Kant's account, because it is the self-consciousness that constitutes me as the numerically identical subject of thinking throughout the manifold of representations given to me in combination in general as a manifold I can relate to an object in and through determining (in reflection) how I am to do so, Kant says that the 'the identity of apperception' 'precedes all *my* determined thinking a priori' (B134).

'Previously,' 'first,' 'precede' – all have here an explanatory, and not a temporal, sense. Indeed, Kant stresses that the 'combination of a manifold in general' in which a manifold of representations must be given to a thinker's consciousness to be combined in the object is as such fundamentally constitutive of all combination of a given manifold in the object. This is why he says combination in general 'must be originally unitary and equally valid for all combination' (B130). And this holds for any combination of a manifold of representations – whether 'we are conscious of the combination or not, whether it is a combination of the manifold of intuition or of several concepts' (ibid)

In the closing sentences of Section 15, Kant sets the task of Section 16 as explaining how, in being added to the representation of the manifold, and to the synthetic unity that combination in general gives this manifold, 'the representation of this synthetic unity' 'first makes the concept of combination possible':

But the concept of combination includes, besides the concept of the manifold and of synthesis, also the concept of the unity of the manifold. Combination is the representation of the *synthetic* unity of the manifold. The representation of this unity cannot, therefore arise from the combination; rather, by being added to the representation of the manifold, it first makes the concept of combination possible. This unity, which precedes all concepts of combination a priori, is not the former category of unity (Section 10); for all categories are grounded on logical functions in judgments, but in these combination, thus

the unity of given concepts, is already thought. The category therefore already presupposes combination. We must therefore seek this unity (as qualitative, Section 12) somewhere higher, namely in that which itself contains the ground of the unity of different concepts in judgments, and hence of the possibility of the understanding, even in its logical use. (B130-1)

What Kant seeks in Section 16 and there identifies as the original synthetic unity of apperception, is this representation of the synthetic unity of the manifold (i.e. of the synthetic unity that combination in general gives to the manifold) that first makes *the concept* of combination possible. And he identifies pure apperception as that in which we are to find this unity, and this as what 'itself contains the ground' even of the unity of different concepts in judgment. Pure apperception, then, is what Kant contends adds the original synthetic unity of apperception to what is given in the 'originally unitary' action of combination in general. And pure apperception's representation of this unity is 'the actus of spontaneity' that, in bringing forth the representation <I think> (B132), 'first makes the concept of combination possible.'

### Principles of thinking, reflection, and intellectual form

Before we turn to the Claim, we need, finally, to fill out somewhat Kant's account of what thinking consists in. I will do so by explaining some more crucial terminology that, despite not occurring in the Claim itself, will prove useful to have in hand in interpreting the Claim. I start with his notion of a principle of thinking.

Kant characterizes principles (*Principien*) as 'cognitions that are grounds of grounds that follow a certain rule' (AA 27: 749). And a rule, in turn, is 'an assertion under a universal condition' (AA 9: 121). The concept <gold> is a rule, in this sense, one that specifies that the marks it contains (<body>, <yellow>, <dense>, <malleable>, <dissolves in *aqua regia*>, etc.) constitute the universal condition of something's being gold. The principle of contradiction is the principle of our analytic cognition in that it is the cognition that is the ground of a discursive mark being the ground of an analytic cognition: this principle is, for example, the ground of the discursive mark <malleable> being, in the analytic cognition <All gold is malleable>, the ground of this cognition that follows the rule <gold>. The force of 'a priori' in 'a priori principle' is to specify that the principle is as such a cognition that makes what is a ground in the objective order of explanation the ground of the cognition and this out of a consciousness of how that ground objectively necessitates its consequence (Smit 2009). An a posteriori principle, by contrast, takes what is, in the objective order of explanation, a consequence, and makes it the ground in the order of one's cognition. Thus Kant tells us that the particular law of nature <All bodies are heavy> serves as an a posteriori principle when one derives from it the

cognition <This body is heavy> (AA 27: 749): the synthetic a posteriori cognition <All bodies are heavy> is the ground of this cognition one has of the individual body one designates with 'this body' as heavy. But our cognition of the particular law of nature <All bodies are heavy> is, as such, a cognition we take from particular successively given appearances as we are acquainted with them in taking them to belong to what Kant terms 'our complete possible experience' (A232/B283): what makes these successively given appearances particular experiences just is, on Kant's account, their belonging, as its parts, to this universal possible human experience. In the cognition <All bodies are heavy>, then, we make what is, in the objective order of explanation, a consequence (namely, actual experiences) the ground of our cognition <This body is heavy>, which is just to say that this cognition is an a posteriori principle. By contrast, purely intellectual principles of thought (*Principien des Denkens*), such as the principle of contradiction, are a species of a priori principles: namely, ones inherent solely in our higher capacity of cognition (as against those that are further determined by the addition of sensible content, such as the transcendental schemata and the principles of pure understanding) that make our acts of thinking first possible. In the case of the principle of contradiction, the possibility is the formal possibility merely of an act of thinking in general. In the case of the categories, the possibility is that of the thinking that constitutes the formal possibility of our experience in general, and thus the formal possibility of a certain act of cognizing a thing. Any principle of thinking is a priori: it is a cognition in which the subject is conscious of how the ground in the objective order of explanation objectively necessitates its consequence.

'Reflection' ('*Überlegung*,' '*reflexio*') is Kant's term for a subject of discursive understanding's consciousness of how the sources of her cognition (understanding and/or sensibility) objectively necessitate how she is, as a person and so in acts of thinking that are imputable to her, to relate given representations to an object (*Object*). Kant thus characterizes reflection as 'the consciousness of the relation [*Verhältnis*] of given representations to our various sources of cognition, through which alone their relation [*Verhältnis*] among themselves can be correctly determined' (A260/B316). A subject's determined thinking is the act in which she aims, in and through reflection, to determine correctly the relation among representations that are given to it in combination in general. Successful thinking realizes form, that is, determination (positing to the exclusion of the opposite), in some matter, that is, the determinable manifold of representations given in combination in general, to determine a thought (A261/B317). So, for example, the logical form of categorical judgment is a form, in this sense, that one realizes in some matter (the concepts <horse> and <mammal>) to determine the relation they have to one another in a categorical judgment ('Horses are mammals'). The principles of our thinking are given to us in

reflection on the form that is determined solely by the nature of our understanding. Let's call such form 'intellectual form'.

A subject's determined thinking consists, not just in reflection, but in that subject's determining its own act of thinking out of its *de se* consciousness of how principles of thinking objectively necessitate that it do so. This follows from Kant's theory of concepts, on which a concept is a rule that a thinker gives to herself in logical consciousness – that is, her consciousness of how the principles of thinking given to her in logical reflection, as intellectual form, determine how she, in her thinking, is to relate representations that are given to her to an object. Here a passage in the *Anthropology* – in which, to look ahead, Kant specifies that logical, or discursive, consciousness is 'pure apperception of one's mental action' – is helpful:

Because experience is empirical cognition, but for cognition (since it rests on judgments) (*reflexio*), and consequently consciousness of the activity [*Tätigkeit*] in the composition [*Zusammenstellung*] of the manifold of representations according to [*nach*] a rule of the unity of the same [i.e. the manifold], that is, concepts and (from intuition distinct) thought in general, is required: so consciousness is divided up into discursive consciousness (which as logical consciousness must lead the way, since it gives the rule), and intuitive consciousness. Discursive consciousness (pure apperception of one's mental action [*Gemüthshandlung*]) is simple. The I of reflection contains no manifold in itself and is always one and the same in every judgment, because it is merely the formal element of consciousness. On the other hand, inner experience contains the material of consciousness and a manifold of empirical inner intuition, the I of apprehension. (7: 141-2)

The intellectual form of a cognition is a purely intellectual principle of thinking that a thinker realizes in some matter (a manifold of representations given to it in combination in general) to determine that matter (the relation among these representations) a priori. One species of such form is the logical form of our thinking in general (which belongs to the subject matter of pure formal logic, as well as transcendental logic). Another is the intellectual form of our pure thinking of a thing (treated only in transcendental logic). The latter, in us, are the categories, concepts that are particularly fundamental species of principles of our thought, that is, 'self-thought a priori first principles of our cognition' (B167).

## The Claim

I want now to sketch a reading of the Claim on which the possibility that it specifies as a necessary condition of a representation being mine is determined purely intellectually, and indeed one that is determined solely by the nature of discursive understanding (of which our higher capacity of cognition is

a species). Appreciating this point will allow us to see that the accompaniment of a representation with the <I think> is what brings forth, and constitutes, the intellectual form of the act of thinking in which a thinker relates that representation to an object. This form is, itself, the determination of an objectively necessary relation that a representation of the sort in question (i.e. a concept, an intuition, or a sensation) must stand in with the other representations with which it is, in synthesis, given successively, if it is to be one that is given to that thinker in such a way as to constitute a representation that has a use for that thinker. It will prove illuminating to start by considering what Kant means by 'Das: **Ich denke**'.

I propose that Kant uses '**Ich denke**' to refer to a subject of discursive understanding's purely intellectual *de se* consciousness of itself merely as the subject of spontaneity (self-activity) in pure thinking itself, and so apart from how, in its thinking, it realizes this activity in any of its sensible representations. This echoes the sense in which Descartes uses '*cogito*.' Indeed, that it does helps to explain why Kant does not, at B131, pause to explain what he means by '**Ich denke**,' despite this being the first appearance of this terminology in the (B-edition) *Critique*. I propose, moreover, that in adopting this terminology Kant means to invoke a neo-Platonic tradition (which stretches through Descartes back to Augustine and Plotinus) on which any thinker can – and must, in order, in the course of first philosophy, to come to cognize how it is properly to employ its mind – withdraw its mind from the senses to isolate its pure understanding.<sup>15</sup> Indeed, he has, in introducing the project of the Transcendental Analytic, clearly signaled that this project is, in this crucial respect, continuous with this tradition: in completing this project, he tells us, 'the pure understanding separates itself completely not only from everything empirical, but even from all sensibility' (A65/B89; cf. A65-6/B90-1).

We are now finally in a position to appreciate the sense and motivation of the Claim.

The **I think** must **be able** to accompany all my representations [*muss alle meine Vorstellungen begleiten können*]; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible, or at least nothing for me. (B131-2)

The Cartesian-Augustinian connotations of 'I think' suggest that the Claim is, itself, concerned only with the most general *purely intellectual* condition of the possibility of thinking the representations that are to be given to me for this thinking in combination in general. The Claim's place in the argument of the B-edition Deduction confirms that his aim in Section 16 is to isolate purely intellectual conditions of this sort: we have seen that this section aims, in particular, to identify the original synthetic unity of apperception, a unity that has pure apperception as its source, as the original

representation of synthetic unity – a contribution of the pure understanding – that, ‘by being added to the representation of the manifold, first makes the concept of combination possible’ (B130). Moreover, the condition that the Claim itself specifies is the *most general* such condition in that it holds for *all* representations that can be given, in combination in general, to a thinker for its use in thinking, regardless even of whether those representations are concepts, intuitions, or sensations. To be sure, one actually accompanies a representation with <I think> only in determining, under some more determined principle of thinking, the relation (*Verhältnis*) it stands in, as the sort of representation it is, to other representations with which it is given to one in combination in general.<sup>16</sup> But the Claim itself prescind from any differences among these principles. As such, it must be distinguished from all other purely intellectual conditions that we can specify in doing logic. And, in specifying the most general purely intellectual condition that *all* representations that are to be given to me in combination in general must satisfy, if they are to have any use in my thinking, the Claim constitutes the principle of thinking under which all other principles of my thinking fall, merely as principles of my thinking.

Even the principle of contradiction is a principle of thinking that, itself, determines the formal logical possibility of *concepts*, and thus one that has a more limited scope than does the Claim. Here it is important to recognize that Kant states this principle as ‘the proposition that no predicate pertains to a thing that contradicts it’ (A151/B190). He thus conceives of this principle, in traditional Aristotelian fashion, as one that concerns things (*Dinge, res*) – that is, subjects of activity and power. And what, on his account, this principle determines is, at bottom, the formal logical possibility of concepts having their proper and defining use as concepts – namely, that in a thinker’s act of relating *sensible intuitions* that can be given to that thinker to a thing in an act of thinking that thing. In this way, Kant holds that the principle of contradiction owes its standing as a principle of our thinking to the standing that the categories have as principles of thinking that determine the purely intellectual possibility of given sensible intuitions having a relation to objects (*Gegenstände*) in our thinking.<sup>17</sup>

On Kant’s account, our higher capacity of cognition itself contains the a priori concepts <concept> <intuition>, <sensible intuition>, <sensation> – indeed all the concepts included in our a priori and purely intellectual concept of a possible experience in general. Our higher capacity even includes a purely intellectual concept <succession>. It must include all these concepts, in order to be able, purely a priori, to prescribe to the manifold of representations that are to be encountered in me their conformity to purely intellectual conditions of their having a relation to an object in a thought of the same. Moreover, because the intellectual conditions that determine the possibility that the Claim asserts to be necessary

are specified in purely intellectual principles of thought, it is only insofar as 'all my representations' fall merely under these purely intellectual concepts <concept>, <intuition>, <sensation>, etc. that these principles set the intellectual conditions of the possibility of these representations being 'anything to me.' The conditions that the distinctive nature of our pure understanding itself sets, conditions that render more determinate the possibility that the Claim asserts, thus, are not only impersonal (in that they hold for possible human subjects other than myself).<sup>18</sup> They hold for any representations of the relevant sort (concept, intuition, or sensation) that can be given to me, merely as transcendental subject, in combination in general, on pain of their not being mine. And they do so while prescinding from all determinate sensible conditions of the possibility of my thought and cognition that are set by my sensibility, whether by its nature or in and through its operation.

Kant uses 'accompany' to convey that my thinking a manifold given to me for thinking in combination in general does not, in any way, alter what is so given, or the giving of it in combination in general. The force of the Claim, then, is that for manifolds of representation that are given to me in combination in general to have any use for me, the thinker, they must conform to the intellectual conditions that our higher capacity of cognition itself sets on their having this use.

Consider now the argument Kant gives for the Claim. Kant advances the Claim as an analytic cognition, one that we can cognize a priori as necessarily true under the principle of contradiction by analyzing the concept <my representation> and bringing to light that it contains the concept <representation that it is possible for the <I think> to accompany>.

The **I think** must **be able** to accompany all my representations [*muss alle meine Vorstellungen begleiten können*]; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible, or at least nothing for me.

In the second clause, Kant gives a brief *reductio* argument for the Claim. Suppose that the <I think> could not accompany one of my representations, <r>. In that case, something would be represented in me – that is, the predicates that make up the representational content of <r> – that could not be thought at all. In the case of a concept, this amounts to its not being a possible representation at all, since a concept has a relation to an object only in thinking. In the case of a sensible intuition, or a sensation, however, all that follows is that it cannot have any use in my thinking, and so is 'nothing for me,' the thinker. It may still have a use in an operation of my lower capacity of cognition, which operation relates sensible intuitions and sensations to objects in and through a merely associative synthesis of the imagination.

On the present reading, the possibility of the <I think> accompanying one of my representations <r> constitutes the purely intellectual possibility of my



being able to think  $\langle r \rangle$ . Moreover, the  $\langle I \text{ think} \rangle$ 's actually accompanying a representation  $\langle r \rangle$  in the synthesis, and so together with the other representations that make up the manifold that is given to me in combination in general, is what constitutes my realizing some intellectual form in this manifold to think  $\langle r \rangle$ , and thereby my putting it to use in relating it to an object in an act of thinking. Contrary to what most readers of Kant have assumed, the  $\langle I \text{ think} \rangle$ 's actually accompanying a representation that is given to me in combination in general is constitutive, in respect of its intellectual form, of my putting it to any use in my determined thinking. This accompaniment is not to be understood on the model of the sort of explicit self-ascription that one might take to be expressed by 'I think p'.

### Pure apperception, original synthetic unity of apperception, and understanding

In the service both of developing the present reading of the Claim and of clarifying his account of pure apperception, and its unity, I want now to sketch, if only briefly, how immediately after stating the Claim Kant goes on to derive this account from the Claim. We will see that the Claim serves, thereby, as the *ratio cognoscendi* of the principle of the original synthetic unity of apperception. But it does so in virtue of an a priori insight, achieved in Section 16, into how my pure apperception is what, in representing the original synthetic unity of the apperception of any manifold of intuition that is to be encountered in me, makes all my determined thinking first possible for me. Pure apperception is itself an actus of spontaneity that realizes the synthetic unity of apperception in all manifold of intuition that can be given to me – that is, the original synthetic unity of apperception – as my capacity of understanding itself (B134n). It is also what, in realizing this unity in a manifold of representations that is given to me in combination in general combines it in my concept of an object in general. Indeed, in doing so, it is what brings forth the representation  $\langle I \text{ think} \rangle$  (B132). But all my determined thinking consists in my accompanying with the  $\langle I \text{ think} \rangle$ , representations that are given to me in combination in general. In this way, pure apperception is the *ratio essendi* of our capacity of thinking itself.

Recall that these purely logical conditions of the possibility of the  $\langle I \text{ think} \rangle$  accompanying my representations differ, depending on whether the representations given to me for thinking are intuitions, sensations, or concepts. Immediately after stating the Claim, Kant at B132 draws our attention to the distinctive necessary relation (*Beziehung*) that *all manifold of intuition* has 'to the I think in the same subject in which this manifold is to be encountered.'

That representation that can be given prior to all thinking is called intuition. Thus all manifold of intuition has a necessary relation [*Beziehung*] to **I think** in the same

subject in which this manifold is to be encountered. But this representation is an act [*actus*] of **spontaneity**, i.e. it cannot be regarded as belonging to sensibility. I call it the **pure apperception**, in order to distinguish it from the **empirical** one, or also the **original apperception**, since it is that self-consciousness which, because it brings forth [*hervorbringt*] the representation **I think**, which must be able to accompany all others, and which in all consciousness is one and the same, cannot be accompanied by any further representation.

The necessary relation in question is, as we have seen, one that this manifold has to the subject where it is considered merely as transcendental subject, and so prescinding from all the sensible content of its thought and cognition. To be sure, any transcendental subject is, as such, a subject that has some distinctive sort of sensibility, one that determines the distinctive forms of its sensibility. Moreover, any thinker constitutes a transcendental subject only on the condition that its sensibility is originally affected, timelessly, by some transcendental object (*Gegenstand*) or other that, in this affection, determines all possible appearances that belong to a single universal possible experience. Nonetheless, the nature of the capacity to think that a subject has merely as transcendental subject does not specify either the distinctive nature of its sensibility, or the manner of the original affection of its sensibility. Thus, the scope of 'all manifold of intuition that is to be encountered in it' encompasses, not only all the pure intuitions contained in the forms of that subject's sensibility, but also all the appearances that are determined by whatever transcendental object (*Gegenstand*) originally affects its sensibility.

A thinker's original representation of the necessary relation (*Beziehung*) that all manifold of intuition that is to be encountered in it has to the <I think> is itself, Kant then tells us, 'an *actus* of spontaneity', which he dubs 'pure apperception' (B132). I propose that, on Kant's account, pure apperception (the self-consciousness of an understanding that is discursive, as against intuitive cf. B136 and B138-9) is the single a-temporal mental act in and through which a thinker prescribes to any and all manifold of intuition that can be given to its consciousness (where the thinker is considered merely as transcendental subject) its objectively necessary conformity to *any and all* of the purely intellectual conditions of the possibility of the <I think> accompanying it. These purely intellectual conditions include most fundamentally, but not exclusively, those specified in its concept of an object in general as the purely intellectual conditions of the possibility of the <I think> accompanying the manifold of a given intuition in thinking that constitutes cognition of the object of that intuition.

To be sure, we do not in pure apperception think the conformity that we prescribe *under this description*. Indeed, in pure apperception itself we are not conscious of any determined principle of thinking (i.e. any principle of thinking other than that which the Claim itself states). But, on Kant's account, pure apperception is a thinker's self-activity that, as transcendental

consciousness, eminently contains all the purely intellectual principles of its thinking: these principles are contained in this activity (intellectual consciousness itself) as the possible ways in which it can, and must, differentiate itself in uniting manifolds that are given to the subject in combination in general, much as the essences of created beings were thought, in traditional transcendental philosophy, to be contained in God's single unchanging and unchangeable pure act of being. For these principles owe their standing as such principles to their being purely intellectual cognitions that our capacity of the absolute spontaneity of representation brings forth itself to actuate itself as our capacity of understanding.<sup>19</sup> And pure apperception itself only occurs as the formal element that a-temporally realizes in the temporally successive manifolds of representations given to a thinker through its inner sense different purely intellectual principles of thinking as the intellectual form that determines their respective distinctive relations to an object in acts of thinking. The acts of thinking that determine this relation employ the categories, the purely intellectual concepts that constitute our concept of an object in general, and this in representing the purely intellectual conditions of representations that are given in our intuition having any relation to an object in our thinking. The distinctive nature of our discursive understanding, in determining the character of our pure apperception, determines the twelve categories Kant displays in his table of the categories as the fundamental purely intellectual conditions to which all intuitions that are to be encountered in me must conform if they are to be anything to me, and thereby as the 'self-thought a priori first principles of our cognition' (B167).

Pure apperception itself is, on Kant's account, to be distinguished from the categories, and more generally from all the principles of thinking that it brings forth. The act of pure apperception itself does not change, even as it realizes different intellectual forms in different sorts of representations to determine different sorts of acts of thinking. For example, in realizing our capacity of understanding in its real use, pure apperception realizes different categories in different particular given manifolds of intuition, to relate them to (perhaps only merely possible) things (i.e. to *Gegenstände* of an experience that is possible for us). But it does so timelessly, and in doing so does not change, any more than God, according to orthodoxy, changes in timelessly creating and sustaining creation. These intellectual forms, on Kant's account, *flow* from pure apperception. Or, to put the point in terminology that Kant avoids because of its unreconstructed metaphysical implications, pure apperception *emanates* these logical forms to realize them in that thinker's particular acts of determined thinking. In this way, as Kant puts it, pure apperception is a 'pure, original, unchanging consciousness' (A107). And, on his account, it must be, because its numerical unity – one that constitutes a thinker as what is numerically identical in the strict Leibnizian sense throughout all

representations that can be given to it, through inner sense, for combination – is what ‘grounds all concepts a priori’ (A107). Pure apperception is, in short, a distant descendent (through Plotinus, Augustine, Aquinas, and Descartes) of Aristotle’s active intellect (*de Anima* III 5). Pure apperception is, however, much more restricted in scope than its predecessors, in that all this *actus* eminently contains are the purely intellectual principles of thinking – the purely intellectual a priori principles of our thought and cognition. Moreover, Kant identifies it only as what our capacity of understanding itself *would have to consist in*, and thus as what we have an a priori entitlement to presuppose, transcendently, as an *actus* of spontaneity in which we actually realize ourselves a subjects of thought and cognition, even though, on his account, we cannot, in any way, prove that such an *actus* is really possible, let alone actual.<sup>20</sup>

B132-3 also expresses Kant’s view that the *de se* consciousness that a thinker has, in pure apperception, of itself as the subject of pure thinking – and this, in virtue of pure apperception’s bringing forth its representation <I think> – is what constitutes that thinker as the numerically identical thinker throughout all its acts of determined thinking. I am proposing, then, that the force of ‘in all consciousness is one and the same’ at B132 is ‘numerically the same.’ I am proposing, moreover, that this consciousness must, on his account, be actually present, in all of that thinker’s thinking, as what constitutes the intellectual consciousness it has as the individual thinker it is. We can appreciate this by attending to how Kant goes on, in B133, to identify the unity of the purely intellectual self-consciousness had in pure apperception – one that is to constitute the ‘thoroughgoing identity of the apperception of a manifold in intuition’ – with the consciousness of *a certain synthesis* of the representations that make up this manifold, a synthesis the consciousness of which makes this identity possible:

Namely, this thoroughgoing identity of the apperception of a manifold given in intuition contains a synthesis of the representations, and is possible only through the consciousness of this synthesis. For the empirical consciousness that accompanies different representations is by itself dispersed and without relation to the identity of the subject. The latter relation therefore does not yet come about by my accompanying each representation with consciousness, but rather by my adding one representation to the other and being conscious of their synthesis. (B133)

This synthesis is what Kant later terms the intellectual synthesis, in distinguishing it from the figurative (B151). Pure apperception just is, on Kant’s account, the *de se* consciousness I must have of what I am doing in this synthesis – that is, prescribing to the manifold I unite in this synthesis its conformity to whatever particular purely intellectual conditions I think in this synthesis. And this *de se* consciousness, as applied in the thinking in which I unite the representations that make up this manifold in one

consciousness, is what gives the representations I unite in this synthesis their relation to myself, as the numerically identical thinker throughout the successive consciousnesses of those representations – even though I am not thereby conscious of these representations as my representations. And the synthetic unity of one's apperception that one realizes throughout one's in-themselves-discrete consciousnesses of each element of the manifold one unites consists in the single actus he has dubbed 'pure apperception.' And as the act that timelessly realizes the original synthetic unity of this apperception itself in all manifolds of intuition that are to be given to me in the synthesis of my imagination, my pure apperception constitutes my capacity of understanding itself: Kant thus tells us that the original synthetic unity of this apperception, considered as capacity (the inner possibility of my power to understand), 'is the understanding itself' (B134n). This parallels how, in the A-edition Deduction, Kant draws on his account of how original apperception sets purely intellectual conditions on a representation 'representing something in me' (A116) to argue that 'the unity of apperception in relation to the synthesis of imagination is the understanding' (A119). The exercise of our capacity of apperception, insofar as it realizes synthetic unity of the apperception of all intuition that is to be given to one in this synthesis, is what constitutes our capacity of understanding itself.

Consider now briefly the sense and motivation of Kant's claim that what makes the representation <I think> the representation of ourselves as the subject of thinking that it purports to be can only be its originating in pure apperception, the actus of spontaneity that realizes our capacity of apperception as our capacity of understanding. Recall that a representation is what *is to be* related to an object. What makes the representation <I think> a representation, then, is its being such that the subject that has this representation is to relate it to itself in thinking itself as the subject that has and exercises the capacity of thinking itself. And there must be an intellectual principle under which it is to do so, that is, a principle that makes the <I think>, as it is thought assertorically by a subject, a *de se* representation of that subject as the subject to whom acts of thinking are to be imputed. This principle is the principle of the original synthetic unity of apperception. The sense of Kant's claim, then, is that pure apperception is, in and through its original synthetic unity, the *ratio essendi* of the representation <I think>.

We can now see how, on Kant's account, even sensations have a use in my thinking, and are anything to me, only in virtue of having a relation to things that we are to determine under a principle of thinking that derives its standing as such a principle from the Claim. The Claim, as applied to all manifold of intuition, yields the principle that this manifold has a use for me only as representations that I am to determine under the categories in cognition of their objects. But for this to be possible, the sensations that provide the matter of the appearances that make up the manifold of a given

empirical intuition must, as such (and so as objective sensations), have a relation to objects in thinking, albeit thinking that does not constitute cognition of these objects. Recall that sensations are subjective perceptions (*perceptiones*) (A320/B376), in that they are to be related to objects of our experience in thinking, but only in thinking that attributes these sensations to these objects in relation to the character of our sensibility; they are merely subjective, and not objective, determinations of these objects. The principle of thought under which sensations have this relation, and through which concepts of objective sensations are possible, is thus a principle of thought that is subordinate to, and derives from, the Claim, via the principle of the synthetic unity of apperception (B135-6).

### The obscurity of pure apperception

The present reading of Section 16 confirms my contention that all my determined thinking, on Kant's account, consists in my actually accompanying representations with the <I think>. On this reading, Kant argues in B132-3 that pure apperception is a thinker's actus of spontaneity in which, in bringing forth the representation <I think>, that thinker is conscious *de se*, and purely intellectually, of itself, as it is the subject of the actus of spontaneity in virtue of which it constitutes an individual thinker that is numerically identical throughout all its thinking. But one might worry that the position I am ascribing to him is so manifestly mistaken that we shouldn't ascribe it to Kant – at least as his considered position. Isn't it obvious that we commonly employ our understanding to form concepts of all sorts of subject-matter without being conscious of our doing so? Indeed, it isn't obvious that we ever are, or even can be, conscious of ourselves, in thinking, as subjects of the activity that constitutes us as numerically identical subjects throughout our thinking.

First, let me note that the present reading finds ample confirmation in the Paralogism chapter. There Kant tells us that the I of <I think> 'occurs in all thinking,' when he characterizes rational psychology as attempting to derive a purely intellectual science of the soul as a rational doctrine that 'independently of all experience can be inferred from this concept I insofar as it occurs in all thinking' (A342/B400). Indeed, he implies that the categories are essentially actually accompanied by <I think> when he describes these concepts, referring to them as 'transcendental concepts,' as ones 'which say "I think substance, cause, etc."' (A343/B401): indeed, this is what he means in describing the categories as '*self-thought* a priori first principles of our cognition' (B167, my italics). A few paragraphs later, he tells us that 'the simple and in content for itself wholly empty I' is 'a mere consciousness that accompanies every concept' (A346/B404). He continues 'Through this I, or He, or It (the thing) which thinks, nothing further is represented than a transcendental

subject of thoughts = x' (ibid). He then, speaking of this consciousness – i.e. pure apperception, or transcendental consciousness, consciousness of oneself as transcendental subject – remarks 'the consciousness in itself is not even a representation distinguishing a particular object [*Object*], but rather a form of representation in general, insofar as it is to be called a cognition [*Erkenntnis*]; for of it alone can I say that through it I think anything' (ibid).

My reply to the worry, briefly, is that it is predicated on a misunderstanding of Kant's position. Kant agrees that we are not, and indeed cannot, be conscious of ourselves as transcendental subjects, *provided 'conscious' is used in the sense that requires the degree requisite for clarity*. On his view, the exercise of the spontaneity that is pure apperception, and that, in realizing the intrinsic unity of this apperception, constitutes (in relation to the synthesis of imagination) our capacity of understanding, is, as such, obscure. It is only the self-consciousness I can, in realizing pure apperception in some particular manifold given to me through my inner sense, have of myself as the subject of *that* manifold – an empirical consciousness of myself – that can be clear. Consider here a stretch of the A-edition Deduction that parallels much of Section 16 of the B-edition Deduction:

All empirical consciousness has, however, a necessary relation to [*Beziehung auf*] a transcendental (preceding all particular experience) consciousness, namely the consciousness of myself, as original apperception. It is therefore absolutely necessary that in my cognition all consciousness belong to one consciousness (of myself). Now here is a synthetic unity of the manifold (of consciousness) that is cognized a priori, and that yields the ground for synthetic a priori propositions concerning pure thinking in exactly the same way that space and time yield such propositions concerning the form of mere intuition. The synthetic proposition that every different empirical consciousness must be combined into a single self-consciousness is the absolutely first and synthetic principle of all our thinking in general. But it should not go unnoticed that the mere representation **I** in relation to all others (the collective unity of which it makes possible) is the transcendental consciousness. Now it does not matter here whether this representation be clear (empirical consciousness) or obscure, even whether it be actual; but the possibility of the logical form of all cognition necessarily rests on the relation [*Verhältnis*] to this apperception [what he here refers to as 'transcendental consciousness' but also as 'pure apperception'] *as a capacity*. (A117n)

The consciousness – pure apperception – relation to which grounds 'the possibility of the logical form of all cognition' is transcendental, and not empirical. But this transcendental consciousness itself, as it represents and prescribes to all manifold of intuition that is to be encountered in one (merely as transcendental subject) its necessary conformity to principles of thinking (as the purely intellectual conditions of their being anything to me), is obscure. And to say that it is obscure just is to say that the *de se* consciousness of myself, as thinker, in which it consists, does not suffice to

provide me with a representation by means of which I can represent myself as the same or different though different representations. It is only as this consciousness takes up some particular manifold given to me in my inner sense that it provides me with grounds for distinguishing myself – and this, only in inner experience – as a subject that is different from other (perhaps merely possible) subjects. So Kant's account of how, in doing transcendental logic, I am to isolate pure apperception, or even the <I think>, had better *not* be one on which I am supposed to draw on a clear self-consciousness of myself as the subject of thinking. Notice that the reading I sketched of Section 16 is one on which I am not.<sup>21</sup>

The recognition that, on Kant's account, pure apperception is itself essentially obscure has direct bearing on Kitcher's proposal that his account of higher cognition is consonant with the agential approach to self-knowledge. On this account, as Kitcher interprets it, judging and inference require conscious acts in such a way that 'whenever a subject is consciously thinking that *p*, she can self-ascribe the activity and the thought' (Kitcher 2017, 171). Now the consciousness that, according to Kant, one must have in consciously believing *p* gives one grounds for self-ascribing the belief *p* only if that consciousness is clear. But, as we have seen, the purely intellectual consciousness that, on Kant's account, one must have, *de se*, of oneself in thinking, as the subject of that thinking, is itself obscure. On his account, it is only the logical consciousness in which one gives oneself the rule for that thinking that is, and need be, clear. In this way, Kant has a principled reason to maintain, as Kitcher herself recognizes that he does, that a thinker self-ascribes representations only in being aware of them through inner sense, and so through a sort of observation of her own mental states.

Moreover, we can see that Kitcher is also mistaken in reading Kant as holding that a thinker, in consciously drawing inferences from her various mental states, thereby creates and recognizes 'the relation of necessary connection across them that makes them the states of a single thinker' (Kitcher 2017, 171). What makes the various mental states that a thinker has, in succession, its states is not the determined thinking in which that thinker relates their representational contents to an object or, for that matter, to one another. It is, rather, the thinker's exercise of its originally unitary actus of combination in general in pure apperception, that, in prescribing to these representations their objectively necessary conformity to the purely intellectual conditions of their being accompanied with <I think>, makes possible their collective unity in that thinker that makes them its representations. The consciousness of this collective unity – the original synthetic unity of the apperception of these representations – is obscure. Nonetheless, it must be present in all my thinking, as the *de se* consciousness that makes my representations mine (gives them use in my determined thinking) in timelessly taking them up into itself.



## The a priori warrant for principles of thinking, and a defense of Kant's account of pure apperception

Some will balk at reading Kant's account of the intellectual conditions of possibility of our thinking as grounded in pure apperception, where pure apperception is understood as I propose: that is, as a single supersensible, and so a-temporal, exercise of the spontaneity of our cognition that, in an Aristotelian vein, constitutes a thinker as the numerically identical subject of all its thinking. For to do so will seem to many to read this account as making a posit that is, by Kant's own lights, far too metaphysically extravagant – one more fitting of a dogmatic rationalist such as Leibniz. For these claims about our capacity of understanding would, it seems, have to consist of theoretical cognition of ourselves as subjects of supersensible acts. How, they will ask, can Kant advance such cognition without violating the central negative tenet of his critical philosophy – namely, that we cannot have any theoretical cognition of things in themselves? I cannot hope to provide a fully adequate reply to this worry here. But let me briefly sketch a line of interpretation that promises to explain how Kant does not, and given his purposes need not, advance his account of pure apperception as a body of cognition of the thing that thinks – or indeed, of any thing at all, let alone a supersensible thing in itself. He need not, because he does not aim to establish that we are subjects of apperception, or indeed that what we conceive of in thinking of pure apperception is even really possible. Kant aims only to establish our a priori warrant to *presuppose* that we are, as thinkers, subjects of pure apperception. And this a priori warrant is, he contends, sufficient for us to have genuine principles of thinking, as we must if we are to be subjects of acts of thinking that are imputable to us. The most fundamental of these principles, as we have seen, are the purely intellectual ones that constitute all the intellectual forms of acts of thinking of which we are capable, given only the nature of our understanding. These include, most crucially, the categories: the purely intellectual concepts through which we can cognize things in respect of the being they have in appearing to us, in prescribing to manifolds of intuition that can be given to us their objectively necessary conformity to these concepts.

On Kant's account, I have the a priori warrant requisite to have genuine principles of thinking only if I, in fact, have a capacity of understanding. And I have this capacity only in and through the two acts of spontaneity in and through which I bring given representations to the original synthetic unity of apperception: combination in general and pure apperception. But Kant sets a further, broadly internalist, condition I must meet if I am to have this a priori warrant: it must, if only in principle, be possible for me to achieve the rational insight that Kant purports to achieve, in Section 16, into our capacity of understanding – namely, an insight into how their originating in pure

apperception is not only necessary but *would*, if actual, be sufficient, for these principles to be the purely intellectual principles of thinking that collectively, under the principle of the original synthetic unity of apperception, constitute the intellectual form of any determined thinking that is to be possible for me. The crucial point to see is that, in achieving this insight, what I *cognize* is only the inner possibility of these principles constituting *my* representations – that is, principles that I am, in my thinking, to relate to manifolds as they are given to me through my inner sense as principles to which these manifolds must conform, if they are to have any use in my thinking. To be sure, on the reading I propose, this insight essentially employs a *representation* of the inner possibility of my being as a thing that thinks. But this insight is not an insight into this inner possibility, and the representations of the spontaneity of our cognition that it employs are not, themselves, cognitions. Nonetheless, this insight suffices to provide an account of an origin in an operation of the spontaneity of our cognition for the purely intellectual principles of my thinking – including the categories – that establishes, entirely a priori, the a priori warrant I have to employ them, a warrant I must have if they are to constitute genuine principles of my thinking.

In short, Kant's account of pure apperception, and the use to which he puts it in the Transcendental Analytic of the first *Critique* is not undermined in any way by his contention that we cannot have any theoretical cognition of the transcendental grounds of the possibility of our thinking that this account specifies. All that Kant's account of pure apperception, and of the transcendental subject of thinking that one constitutes in this *actus*, need do is to provide us a mere representation of these grounds that each of us can, merely in virtue of the purely intellectual principles of thinking contained in the nature of our capacity of understanding itself, employ in an act of self-understanding to establish, *de se*, her own a priori warrant, merely as a thinker, to prescribe to all manifolds of intuition that can be given to her in intuition their objectively necessary conformity to the categories. In establishing this in principle possibility, Kant contends, his transcendental deduction of the categories establishes, for each of us, a *de se* a priori warrant, merely as a subject that takes herself to be capable of acts of thinking that are imputable to her, to presuppose, as she does in any of her determined thinking, that the conditions of the possibility of her, in fact, being such a subject actually obtain. These conditions include that she be the subject of an act of pure apperception that realizes a capacity of understanding in her, and thereby the transcendental subject of all her determined thinking. And Kant succeeds in establishing our a priori warrant to make this presupposition by employing a representation of pure apperception merely as an act that *would have* to take place in us, if we are to be capable of any acts of thinking that are imputable to us.<sup>22</sup>

## Conclusion

We have seen that Kant works with a conception of what it is to be a thinker that is importantly similar to the account of what it is to be a critical reasoner that informs the agential approach to self-knowledge. For, on the reading offered here, the Claim asserts that, unless a representation that is given to a subject of thinking meets the condition that the Claim specifies as necessary to its being a representation that this subject can put to use in thinking that is imputable to her, that representation ‘cannot be thought at all’ and is ‘nothing to her.’ In this way, Kant’s Claim expresses an idea that parallels, in an important respect, that which animates the agential approach to self-knowledge: a representation that belongs to me, the person, is as such a representation that can play a proper role in the activity that characterizes me as a person (i.e. a subject that is responsive to reasons in the relevant way).

Nonetheless, we have also seen that what lies behind the Claim is a conception of what it is for us to be thinkers that is, in other respects, importantly different from that which animates the agential approach to self-knowledge, one on which we are, as thinkers, subjects of rational belief. What animates Kant’s position is, rather, a conception on which what makes us thinkers is our having the authority, as subjects of a capacity of conceptual understanding, to prescribe a priori to manifolds of representations as they are given to our consciousness their conformity to purely intellectual principles of our thinking, inherent in this capacity, that determine conditions of these representations having any use in thinking that is to be imputable to us. Moreover, this a priori warrant that, on Kant’s account, we have merely as thinkers is not one to self-knowledge at all. It is, rather, one that each of us, as a thinker, has to take representations as they are actually given to her consciousness as subject to the purely intellectual conditions that our capacity of understanding sets on their having a use in her thinking – as we actually take them to in putting them to use in our thinking.

The reading I propose of the sense and motivation of the Claim has the virtue of helping to clarify the role it plays in the argument of the B-edition Transcendental Deduction of the Categories. We have, if only in outline, seen how on this reading the Claim provides the starting point of Kant’s attempt, in Section 16 of this Deduction, to achieve a certain rational insight: cognition of how and why the essence of the capacity of thinking itself – the principle of the synthetic unity of apperception – is as such the *ratio essendi* of all principles of our thinking; it is only in standing under this, the highest principle of all our cognition, that these principles constitute genuine priori cognitions of the objectively necessary unities that manifolds of representations given to us must have if they are to have any use in our thinking. This insight yields an account of what would have to be the case if we are to have the a priori warrant requisite to our being subjects of acts of thinking that are to be

imputable to us. This points the way to a new approach to the Deduction, one on which it does not depend on an ultimately undefended assumption that we are, in fact, subjects of experience, or even that we are, in fact, subjects of acts of thinking that are imputable to us. The Deduction, rather, aims to achieve a purely intellectual insight into how having their origin in pure apperception would give what we take to be the purely intellectual principles of our thinking the normative ground they require to constitute genuine principles of our thinking. Moreover, this insight serves to establish, entirely a priori, an a priori warrant we have, merely as subjects of discursive understanding, to presuppose that the conditions of the possibility of our being such subjects actually obtain. And this includes an a priori warrant we have, merely as subjects of thinking that is imputable to us, to take manifolds of intuitions that are, and can be, given to us as subject to the most fundamental purely intellectual principles of our thinking, namely, the categories. It is in this way that Kant's Transcendental Deduction of the Categories specifies an origin for the categories that suffices to establish their a priori objective validity, if only for objects of our experience.<sup>23</sup>

## Notes

1. A proponent of the agential approach to self-knowledge need not hold that the only warrant we have, or could have, for self-knowledge derives from our beliefs being embedded in our critical reasoning.
2. The most important such proponent is Tyler Burge, who in a series of influential papers has developed detailed accounts of a priori entitlements that we have as persons, entitlements that include one to self-knowledge. These accounts, as he puts it, 'have a Kantian flavor' (Burge 2003, p. 335 n50). In particular, in 'Our Entitlement to Self-Knowledge,' Burge develops an account of an epistemic entitlement to self-knowledge that turns crucially on the claims that being a critical reasoner requires that one be the subject of 'mental acts and states that are knowledgeably reviewable' (Burge 1996, 98) and that this reviewability requires that one have a certain non-observational entitlement to first-person judgments regarding those acts and states (ibid). Burge's emphasis on the possibility of a subject's applying 'the I think' to her thoughts echoes Kant's famous discussion at B132f. I hasten to add that Burge is careful to distinguish his position from Kant's (see, e.g. Burge 1996, p. 99, n5), and his aim in these papers (Burge 1988, 1993, 1996, 2003, 2011) is not Kant exegesis (although he does offer several penetrating observations about Kant along the way). And I need especially to stress that the reading of B132f., and more generally of Kant's account of the unity of apperception, that I will be rejecting is not one that Burge endorses.
3. See Evans (1982), Boyle (2009), and Setiya (2011).
4. Here 'understanding' is to be taken in the sense in which understanding contrasts with judgment and reason. Kant also occasionally uses this term in a broader sense to refer to our higher capacity of cognition, which encompasses judgment and reason, as well as understanding in the narrower sense of the term. To look ahead, to understand something, in turn, is to represent

its essence positively with a warrant sufficient for that representation to serve as a principle of our cognition. This does not require that the representation constitute cognition of the essence.

5. The term 'transparent' is one coined by Roy Edgley (1969, p. 90).
6. As Patricia Kitcher points out, in reading Kant this way, Evans is following Strawson (see her Kitcher 2017, discussed below).
7. Moran and Boyle are happy to allow that a critical reasoner may have other modes of access to her beliefs in addition to the agential one. All they claim is that agential access is privileged, in that it tracks the proper functioning of the capacity for belief. So the thesis of the transparency of belief to which they subscribe is weaker than that which Evans champions: they claim only that there are some cases in which a critical reasoner can determine what she believes by addressing world-directed questions, and not that the only way in which she can determine what she believes is by addressing such questions.
8. Neither Burge nor Moran reads Kant as advancing, at B131-2 or elsewhere, the claim that our beliefs are transparent.
9. One might worry here that what Kant says in this passage is inconsistent with his famous claim that concepts and intuitions constitute cognitions only in conjunction with one another: this claim seems to entail that concepts and intuitions cannot be species of cognition. But this inconsistency is only apparent. On Kant's account, a concept can be considered either of itself, or as it constitutes a cognition in conjunction with intuition, and the same goes for intuition. At A320/B372, Kant is using 'concept' and 'intuition' to refer to these representations as they constitute cognitions.
10. Burge and Moran, as well as other proponents of the agential approach, maintain that our knowledge of our sensations is of a different character than our knowledge of our thoughts and attitudes. And they exclude sensations from the nonobservational self-knowledge that a critical reasoner can, as such a reasoner, have of her mental acts and states, on the grounds that sensations are not essentially the content of her thoughts and attitudes. See Burge (1996, 107), Moran (2001, 9–10), as well as Aaron Zimmerman (2008) and Matthew Boyle (2009). For a helpful discussion of this point, as well as of what, following Burge, has come to be referred to as the rationalist approach to self-knowledge, see Brie Gertler (2011, especially Chapter 6).
11. A related problem is worth mentioning: Kant distinguishes between two sorts of cognitions – concept and intuition – and it isn't at all clear that to have a *perceptio* that is an intuition is, itself, to think it, or indeed to take any attitude toward it. Kitcher addresses this concern in the course of developing a reading of Kant on which his position is amenable to the agential approach to self-knowledge (Kitcher 2011, 151f).
12. This line of thought seems to raise difficulties only for Evans's transparency thesis, on which the question of what one believes always can, and is to be, settled solely by answering the world-directed question. And, even so, it isn't entirely clear what the difficulty is supposed to be. Kitcher seems to be assuming that the prior *de se* consciousness one has of oneself as the subject of rational thought that Kant's account posits is either itself a belief that cannot be determined in the world-directed fashion, or a source of such beliefs. Thanks to Thomas Land for raising this worry.
13. This is the central thesis of her (2017). See also Chapter 15 of (2011).

14. 'Aktus' and 'actus', in Kant's terminology, signifies, not just any action (*Handlung*), but an actuation, or *inner* action, i.e., a subject's activity (*Activität*) insofar as it suffices to realize its capacity (*Vermögen*) as a power (*Kraft*). See here the final section of my (2009).
15. Where Kant parts company with Descartes, and others, is in denying that <I think> is *cognition of the thing* that thinks: it is a *mere representation of the thing that thinks*, because we cannot prove that any such thing is really possible. He maintains that the representation <I think> has the function of making it possible for one to determine one's own thinking out of one's *de se* consciousness of how one is, in bringing them under intellectual principles, to relate representation to an object. And he holds that the <I think> can serve this function without being cognition of the thing that thinks. It can, and does, serve this function in virtue of constituting cognition that a human being can have, through pure apperception, of herself in respect of 'actions and inner determinations' of which she must, timelessly, be the subject, if she is, as a human being, to be a person at all (A546/B574).
16. Consider here how, in his 4 December 1792 letter to J.S. Beck, Kant reacts to Beck's attempt to gloss the Claim. Kant offers the following emendation to Beck's attempt: 'Instead of ... "The I think must accompany all the representations in the synthesis," "must be able to accompany [*begleiten können*]"' (11: 395). Notice that Kant does not correct Beck's gloss of the Claim in any way other than to add the modal qualifier '*können*.'
17. One might object that the objective validity of the principle of contradiction cannot depend on our thought's relation to things because pure general logic abstracts 'from all content of our cognition, i.e. from any relation [*Beziehung*] of it to the object [*Object*]' (A55/B79). To be sure, the principle of contradiction does, on Kant's account, abstract from all relation of our cognition to the object. But this is entirely consistent, on his view, with its owing its standing as cognition to its being applicable to objects of our experience, much as pure mathematics owes its objective validity to there being 'things that can be presented to us only in accordance with the form of our sensible intuition' (B147), despite its abstracting from the dynamical character of these things.
18. For a helpful discussion of this point, see Keller (1998).
19. For a reading of Kant's conception of the spontaneity of cognition along these lines, see my (2009).
20. Here it is useful to contrast the present reading with a similar one that Dieter Henrich considers, and rejects, in (Henrich 1976), p. 79f. On this reading, Kant conceives of the subject as having strict Leibnizian identity (one that requires that identicals be indiscernible). Henrich rejects this reading as incompatible with Kant's characterization of the capacity that is essential to the subject as one for altering its condition in and through reflection on how it is to think. The present reading provides the resources for distinguishing between pure apperception itself, which constitutes the thinker as transcendental subject, and the alterations it determines in itself as empirical subject.
21. This provides some indication of how the present reading is not subject to a second worry that leads Henrich to reject his reading of the subject as enjoying strict Leibnizian identity (Henrich 1976, p. 79f.). The worry is, in effect, that Kant cannot provide a case for such an account of the subject because we are not given to ourselves, in consciousness, as transcendental subjects. The

response is that Kant does not purport to arrive at his account from what is given to him, or us, in a *clear* consciousness.

22. For details, see 'Essence, Nature, and the Possibility of Metaphysics,' forthcoming. Here I can, given the constraints of this paper, only gesture at the approach to Kant's critical philosophy that I develop and defend here, and elsewhere.
23. This paper was written for the conference *Transparency and Apperception* held at Ryerson University in May 2018. I thank the organizers of this conference – David Hunter, Thomas Land, and Boris Hennig – and other conference participants for helpful discussion. Part of this paper was presented at the UCLA Kant conference held in February 2019. I thank John Carriero for organizing this conference, and to all its participants for their helpful comments. I also thank Robert M. Adams, Tyler Burge, Tom Christiano, Suzanne Dovi, Frode Kjosavik, Thomas Land, Shaun Nichols, Santi Sanchez, and Mark Timmons for helpful comments and discussion.

## Notes on contributor

*Houston Smit* is Associate Professor of Philosophy at the University of Arizona.

## References

- Boyle, M. 2009. "Two Kinds of Self-Knowledge." *Philosophy and Phenomenological Research* 78: 133–164. doi:10.1111/phpr.2008.78.issue-1.
- Burge, T. 1988. "Individualism and Self-Knowledge." *Journal of Philosophy* 85: 649–663. doi:10.5840/jphil1988851112.
- Burge, T. 1993. "Content Preservation." *The Philosophical Review* 103: 457–488.
- Burge, T. 1996. "Our Entitlement to Self-Knowledge." *Part 1 Proceedings of the Aristotelian Society* 96: 91–116. doi:10.1093/aristotelian/96.1.91.
- Burge, T. 2003. "Perceptual Entitlement." *Philosophy and Phenomenological Research* 67: 503–548. doi:10.1111/phpr.2003.67.issue-3.
- Burge, T. 2011. "Self and Self-Understanding." *Journal of Philosophy* 108 (6/7): 287–383. doi:10.5840/jphil20111086/715.
- Edgley, R. 1969. *Reason in Theory and Practice*. London: Hutchinson.
- Evans, G. 1982. *The Varieties of Reference*, edited by J. McDowell. Oxford: Oxford University Press.
- Gertler, B. 2011. *Self-Knowledge*. New York, NY: Routledge.
- Henrich, D. (1976) *Identität Und Objectivität*. Heidelberg: C. Winter Universitätsverlag.
- Keller, P. 1998. *Kant and the Demands of Self-Consciousness*. New York: Cambridge University Press.
- Kitcher, P. 2011. *Kant's Thinker*. New York, NY: Oxford University Press.
- Kitcher, P. 2017. "A Kantian Critique of Transparency." In *Kant and the Philosophy of Mind: Perception, Reason, and the Self*, edited by Anil Gomes and Andrew Stephenson, 158–172. Oxford: Oxford University Press.
- Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.
- Setiya, K. 2011. "Knowledge of Intention." In *Essays on Anscombe's Intention*, edited by Anton Ford, Jennifer Hornsby, and Frederick Stoutland, 170–197. Cambridge, MA: Harvard University Press.

- Smit, H. 2009. "Kant on Apriority and the Spontaneity of Cognition." In *Metaphysics and the Good: Themes from the Philosophy of Robert Merrihew Adams*, edited by Samuel Newlands and Larry M. Jorgensen, 188–251. New York, NY: Oxford University Press.
- Zimmerman, A. 2008. "Self-Knowledge: Rationalism Vs. Empiricism." *Philosophy Compass* 3 (2): 325–352. doi:[10.1111/j.1747-9991.2008.00125.x](https://doi.org/10.1111/j.1747-9991.2008.00125.x).