

# A Kantian Critique of Current Approaches to Self-Knowledge 1: Anscombe's Thought Experiment

Patricia Kitcher

Columbia University

## 1. *Three Current Assumptions about Self-Knowledge*

One theme of contemporary discussions of self-knowledge is that Descartes and Locke offered attractive, but false pictures of the phenomenon. On the other hand, Kant's views were minimalist and correct: They agree with current approaches. I'm going to argue that this evaluation is only half right. Kant's investigation of the necessary conditions for cognition gave him a far better understanding of self-knowledge than his predecessors. His insights into the topic are, however, not in perfect agreement with current approaches. If we look at the theory of the transcendental unity of apperception, then we can see that three foundational assumptions of current work on self-knowledge are mistaken.

The idea of turning to Kant for insights about self-knowledge comes from the work of P. F. Strawson. Strawson presents his essay 'Persons' in terms of finding a middle way between the Cartesian view that

when we speak of a person we are really speaking of two distinct substances...each of which has its own appropriate types of states and properties (1959, 94)

and the no-self view that states of consciousness do not belong to anything. He saw Kant as more perspicuous than Hume because he recognized that

the unity of the 'I think' that accompanies all my perceptions and therefore might just as well accompany none ...is a purely formal ('analytic') unity. (1959, 103)

Strawson's way of going between the bad alternatives is to work out the implications of what he took to be two obvious facts:

It is a necessary condition of one's ascribing states of consciousness, experiences, to oneself, in the way one does, that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself.

The method of verification [is] so different in the two cases [the first person case and the second/third person case]—or rather ... there [is] a method of verification in one case (the case of others) and not, properly speaking in the other case (the case of oneself). (1959, 99, 99-100)

These two claims are presented in his clear account of his central argument:

One can ascribe states of consciousness to oneself only if one can ascribe them to others. One can ascribe them to others only if one can identify other subjects of experience. And one cannot identify others if one can identify them only as subjects of experience, possessors of states of consciousness. (1959,100).

From this argument, it is only a step to Strawson's conclusion that 'person' is a basic term, a term which must be used with both material and psychological predicates. The argument has been criticized, with the most important objections directed at the verificationist claim that there must be 'logically adequate' behavioral criteria for ascribing psychological predicates (1959, 106). Although few philosophers still accept the argument of 'Persons,' many assume that Strawson identified a crucial requirement for any theory of self-knowledge: It must deal with the very different ways that humans have of acquiring [not 'verifying'] self-knowledge and knowledge of other minds. In the last talk, I will argue that, from a Kantian perspective, this assumption about the relation between two kinds of psychological knowledge is backwards. What is distinctive about knowledge of 'other minds' is that subjects can understand others as cognitive subjects *only* through their understanding of themselves as such subjects.

Gareth Evans offered what he described as a preliminary account of the self-knowledge of beliefs (1982, 255). Evans's view has, however, become widely accepted. One reason is that it seems to provide a model of self-knowledge that avoid[s] the idea of this kind of self-knowledge as a form of perception—mysterious in being incapable of delivering false results (1982, 225).

Evans argues that we are led to the Cartesian position by the assumption that we know our minds through taking an inward glance at our states and activities. That is the mistake he tries to correct:

In making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?' I get myself

in a position to answer the question whether I believe that  $p$  by putting into operation whatever procedure I have for answering the question whether  $p$ ...

We can encapsulate this procedure for answering questions about what one believes in the following simple rule: whenever you are in a position to assert that  $p$ , you are *ipso facto* in a position to assert 'I believe that  $p$ '. (1982, 225-26)

Evans goes on to say that other elements must be involved in possessing the concept of 'belief,' such as understanding that the psychological concept ' $\xi$  believes that  $p$ ' can also be instantiated by others (1982, 226)—an obvious reference to Strawson's view.

Among the many philosophers influenced by Evans Richard Moran (2001) takes the view that the inner mental world is 'transparent' to the outer world as a starting place for his account of the authority of self-ascriptions of belief. In the last talk, I will argue that Kant's theory implies that it is not possible to ascribe self-knowledge solely through looking outward.

The third current assumption that is undermined by Kant's work is not usually highlighted. It is implicit in the discussions of Strawson and Evans, but much more obvious in Elizabeth Anscombe's seminal discussion of 'The First Person.' I consider it in that context. Anscombe's focus is the odd situation of 'I-thoughts.' The difficulty is that any solution to the problem of the reference of 'I' must be such that reference is guaranteed (1975, 56). In that case, however, the referent cannot be the body, so it could only be the Cartesian ego.

Anscombe argues for this claim by introducing a *Gedanken*-experiment about a sensory deprivation tank. In the tank Anscombe would be without sight, unable to speak, unable to touch her body, and locally anaesthetized. Under these circumstances the object of 'I' could not be her body, because that object is in no way present to her.<sup>i</sup> Yet she thinks that she might well tell herself:

'I won't let this happen again.' (1975, 58)

That is, she believes that she would not have lost her self-consciousness nor what she means by 'I.' In the sensory deprivation tank, however, there is nothing that could be present to her but a Cartesian ego, *viz.*, 'the thinking that thinks this thought' (1975, 58).

Thus we discover that, if 'I' is a referring expression, then Descartes was right about what the referent was. His position has, however, the intolerable difficulty of requiring an identification of the same referent in different 'I'-thoughts (1975, 58).

Even though a person in the tank could allegedly attach 'I' to a current thinking, she could never be in a position to attach the same 'I' to different thinkings. Anscombe takes this to be a *reductio* of the Cartesian approach and concludes that the only solution is to admit that 'I' does not refer (1975, 59-60).

This conclusion has not been widely accepted. On the other hand, many discussions of self-reference start with Anscombe's thought experiment. Anscombe believes that philosophers should try to find a sense for 'I' before turning to the problem of its referent. But in the experiment, she assumes that her readers will agree that she can use 'I' in the sensory deprivation tank—can use it to refer (1975, 58)—even though the issue of the identity of the 'I' through time has not been considered. This is the widely held, implicit assumption that is explicit in the *Gedanken*-experiment: It is possible to investigate the referring use of 'I' without considering the unity of the thought in question with other thoughts. From Kant's perspective, if 'I' or 'I-think' can legitimately be used, then the thought must already be understood as united with others in a single self.

To sum up, many contemporary philosophical theories of self-knowledge make three assumptions:

1. First person-third person asymmetry: Self-knowledge of mental states such as beliefs is fundamentally different from knowledge of the beliefs of others.
2. Transparency: Self-knowledge of beliefs can be explained solely through beliefs about the 'external' world.
3. Self-consciousness without unity of self-consciousness: The self-ascription of belief can be explained without considering the unity of that belief with others in a common consciousness.

In this talk I'm going to look again at the Kantian position and try to defend it from some obvious criticisms. Then I will argue that Kant's theory shows that Anscombe's famous experiment is incoherent. In the last talk I'll offer a Kantian critique of the transparency approach and of the claim for first person/third person asymmetry.

## 2. *Kant's Account of the Unity of Self-Consciousness Defended (a little)*

To see the inconsistency between Kant's views and these contemporary dogmas, we need to review of his theory of the transcendental unity of self-consciousness (or apperception). It is a crucial fact about the theory that it is introduced in the context of explaining the possibility of cognition and that the cognition at issue is conceptual cognition—cognition where the subject uses marks/concepts that she understands as such, as grounds of cognition. An ox might differentiate his stall because of its door, but he could not know that that is the ground of his cognition (2.59-60). The ability to recognize marks as such is sophisticated, because it involves what contemporary theorists call 'metacognition.' A cognizer who recognizes concepts as grounds of cognition does not just make judgments, she also knows something about them, namely, that she uses a representation that is common to many things, or that has other marks as its grounds of cognition.

As I argued in the first talk, Kant's key move is to connect conceptual cognition to the unity of apperception or self-consciousness. He takes the concept of a cognitive subject to be that of a subject whose states necessarily belong together (A107, A116, A124, B132, B137). Through engaging in cognition a subject makes necessary connections across her mental states that she recognizes as such. For example, she applies the concept 'nine' through partial representations. She recognizes the partial representations, 'one,' *etc.* as the grounds of her cognition 'nine' and so the representation 'nine' as impossible without those partial representations. She also understands partial representations as such, as standing for features that are common to many things. In being aware of a partial representation 'one,' for example, as 'uniting' a current sensory representation with other representations that present a common property, she is also aware that her partial representation would not exist without the sensory representation. As she makes judgments or applies concepts, she thus makes and recognizes necessary connections among representations. In Kant's words, she 'brings' representations 'under' or 'to' the unity of apperception (A108, B134-35, B136-37).

To offer a Kantian critique, I need to show that his analysis of cognition and his claim about the relation between cognition and the unity of self-consciousness are correct or at least plausible. That task is difficult because his position is not minimalist.

It is not all the same to Kant whether we say that the 'I-think' must accompany all my representations or none. Conceptual cognition would be impossible without a relation of recognized necessary connection across mental states (although ordinary cognizers would not express what they understand in these philosophical terms). Further, conceptual cognition is not widely believed to require conscious acts of combining or synthesizing. But the theory of the transcendental unity of apperception asserts that any case of conceptual cognition involves the unity of self-consciousness and that the unity of self-consciousness is possible only through conscious acts of uniting some representations in others. I will try to show that the theory is sufficiently plausible that it can be used to criticize other theories by raising and rebutting two serious objections. I offer further arguments for it when using it to criticize the three assumptions of contemporary theories of self-knowledge.

A common aim of contemporary work on self-knowledge is to avoid questionable mental processes, so going back to Kant's theory of conscious combining may seem to be a big step in the wrong direction. To see that conscious combining is both a familiar phenomenon and necessary for cognition, let's return to the case I began with, the case of inference.

Kant takes all of 'higher' cognition, inferences as well as 'immediate' judgments through concepts, to involve acts. Although he maintains that humans distinguish what they know directly from what they infer (A303/B359), he notes that:

Weil wir des Schließens beständig bedürfen und es dadurch endlich ganz gewohnt werden, so bemerken wir zuletzt diesen Unterschied nicht mehr und halten oft, wie bei dem sogenannten Betrüge der Sinne, etwas für unmittelbar wahrgenommen, was wir doch nur geschlossen haben.

Despite these deceptions, every case of correct inferred cognition involves an act of inference.

Bei jenem Schlusse is **ein** Satz, der zum Grunde liegt, und **ein** anderer, nämlich der Folgerung, die aus jenem gezogen wird, und endlich die Schlußfolge (Konsequenz), nach welcher die Wahrheit des letzteren unausbleiblich mit der Wahrheit des ersteren verknüpft ist. (A303/B359-60).

Kant's picture of inferring is similar to his account of judging. The faculty of reason draws the conclusion from the premise(s) through an inference according to which the truth of the conclusion is absolutely connected to the truth of the premise(s). He knows that inferences can be mistaken. In fallacies, the truth of a conclusion is connected to the truth of premises even though there is no logical relation between the two. But in the non-deceptive, non-fallacious normal case, the act of inferring connects the truth of the conclusion to the truth of the premises and informs the cognizer that her knowledge is inferred. Through this act, she recognizes her inferred or mediate cognition as such—as cognizing one thing through the cognition of other things.

Kant does not discuss inference in the transcendental deduction in part because inference presupposes judging which already implies the unity of apperception. I appeal to inferring, because the mental act is more noticeable than in cases of judging. Consider the following inference:

If the economy is bad, then the party in power will lose.

The economy is bad.

Therefore, the party in power (will) lose.

Nothing strange is going on here. After reading the premises, normal humans draw the conclusion before seeing or hearing it. And they are conscious in the act of drawing the conclusion, as they must be, if they can understand inference as such, as cognition where they know one thing through knowing others. In the last talk, I'll return to the case of inference and argue that the act-consciousness involved is not merely part of the phenomenology of inferring. It is necessary for inferring.

As mentioned in the first talk, although the idea that people consciously perform cognitive acts was widely accepted in the Early Modern period, mental act awareness has only recently returned to philosophical discussions. When considered it is often limited to unusual cases.<sup>ii</sup> So, for example, Alfred Mele offers the example of trying to think of seven animals whose names begin with G (2009, 17ff.). Tyler Burge argues that to have the first-person concept, one must be capable of critical reasoning and

to be capable of critical reasoning, and to be subject to certain rational norms necessarily associated with such reasoning, some mental acts and states must be knowledgeably reviewable (1996, 98).

Anyone who can use the first person concept and can engage in rational cognition must also be capable of—occasionally—engaging in second-order thought about reasons and mental actions. My appeal to the somewhat artificial case of a formal syllogism may also suggest that mental act awareness is unusual.

As we have just seen, Kant thought that the acts involved in inferences were so common that they were sometimes overlooked. As we saw before, he argues that they are necessary for cognizing through concepts. Thus all cases of 'higher' cognition involve conscious acts: Either the cognizer knows something through concepts and she must understand concepts/marks as such, as grounds of cognition, or she knows something through inference and she must understand inferred cognition as such, as based on other cognitions. In both cases, she must be conscious of her mental activity. We can get a better sense of the ubiquity of mental act awareness by looking at three ordinary activities, driving home from work, looking for reading glasses, and shopping for Christmas presents.

Buying Christmas presents often requires serious amounts of reflection, but my concern is not with judgments about what to buy, but with judgments about what things cost. The shopper who judges that the Lego helicopter is €12 has some sensory basis for using this concept and he knows what that basis is even if he does not reflect on it: He saw the euro symbol and the numerals on a sign, the numerals came up when he ran the bar code under a scanner, he saw the symbol and the numerals in a newspaper ad, etc. The mental movement may be less noticeable than in the case of inference, but he is still aware of applying the concept on the basis of, e.g., seeing the sign, and so of using a word that applies to this case and to many others.

The person who is looking for his reading glasses upstairs knows why he is doing so. He knows that he had the glasses after he came home and that he has already looked for them downstairs. In the case of inference, it is not just a matter of, e.g., remembering that he has already searched downstairs; he understands the relation



between that cognition and his current looking around upstairs. Now consider an example that is mentioned in some discussions of consciousness: driving home while your attention is completely taken up in thinking about something else. This experience can be upsetting, since it is puzzling how such a complicated task could be accomplished 'unconsciously.'

My point is that the conscious acts of inferring when trying to find something or in judging the price of a possible Christmas present or (normally) in judging that a traffic light has turned red are not unusual. They are not reflective states; they happen constantly in everyday 'higher' cognition. Sometimes we do not think about the world around us, either because we are thinking about other things, as in the distracted driving case, or simply because we are day-dreaming. But for most people such states are not common. We go about our daily lives using sensory clues to apply concepts and making inferences from what we learn. That is how we are able to understand the world around us and to figure out what we want to do—for example, to buy the Lego helicopter because it costs less than the Storm Trooper helmet.

My defense of Kant's appeal to the mental act awareness involved in higher cognition is limited: It is not a philosophical invention, but a familiar phenomenon that is easily recognized once attention is drawn to it. It is necessary for normal cases of concept use and inference. I have not provided a theory of this awareness. I treat it, as Kant does, as part of the analysis of conceptual cognition, rather than as something to be explained by theories of mind. In using his theory to criticize others, I will bring out some aspects of mental act awareness, but these observations fall short of a full account. I will, however, address an obvious objection from the inner sense debates. Locke's hypothesis of an 'internal sense' is usually regarded as incoherent, because the sense is supposed to be both perceptual and infallible. To lay the ghost of a magical kind of consciousness to rest, I will clarify the role of mental act awareness in self-knowledge of beliefs or judgments.

Tyler Burge has recently argued many historical and contemporary epistemologists have failed to distinguish between philosophical theories of the necessary conditions for psychological states to have objective reference and the

psychological capacities involved in subjects having the states (2010, 12-22). Burge's focus is on perception. He takes Kant's focus to be the sophisticated sort of cognition required for science (2010, 155, n. 4, 156). As is obvious in my examples, I think Kantian higher cognition has a far wider scope. Although Burge's target is perception, his fundamental criticism may seem also to apply to Kant. The objection is that philosophical accounts of objective reference 'hyper-intellectualize' (2010, 12). In Kant's case, the charge would be that his claim that the unity that is necessary for object concepts is the unity of apperception (talk one, p. 14) is an intellectualist myth. According to Kant conceptual cognizers must possess the unity of apperception—they must create and recognize relations of necessary connection across their states. The objection would be that even if philosophers can show that the concept of conceptual cognition presupposes the concept of a unified subject, that philosophical theory should not be offered as an account of the psychological processes or capacities of actual cognizers. Making this move would be just the sort of 'armchair argument' about the relations among psychological abilities that Burge objects to (2010, 209).

Kant's theoretical interests were not psychological, but philosophical. Through the counting example, he investigates what the phenomenon of conceptual cognition involves: uniting a sensory representation with others as indicators of a common property represented by a mark, and/or combining partial representations in a resultant representation where the former representations are regarded as the grounds of cognition of the complex representation. The analysis of conceptual cognition leads to the philosophical claim that it is necessary for conceptual cognition that different mental states are recognized by cognizers as standing in relations that philosophers describe as relations of necessary connection. He argues that conceptual cognition must involve the production and the recognition of the relation of necessary connection across mental states by raising the possibility of a case where there is no act-consciousness. This is the counting case if there is no conscious synthesis. He notes that it would fail to be 'higher' cognition, because the marks/concepts would not be considered as grounds of cognition. Then he examines the relation between judging (under a concept) and the philosophical problem of the unity of the mind. In judging, cognizing subjects combine representations in another representation thereby making a necessary

connection across the states that they recognize as such. The phenomenon of conceptual cognition thus inevitably involves the unity of self-consciousness.

Kant's theory of apperception is minimal in one respect: He does not have a metaphysically loaded account of the unity of consciousness. He takes the 'unity of self-consciousness' to indicate only relations of necessary connection across mental states. Given this minimal conception, he can move from his argument that conceptual cognition involves the production and recognition of relations across mental states that philosophers call the relation of 'necessary connection' to his philosophical thesis that the unity of self-consciousness or apperception is necessary for such cognition. Thus, when he presents the faculty of conscious combining as a special faculty—'transcendental apperception' (A94-95, A106-106, B132, B134n.)—he is not hypothesizing a novel capacity, but highlighting the philosophical importance of his analysis of conceptual cognition. His analysis shows that conceptual cognition requires some faculty or faculties that consciously combine some states in others and recognize the relation of dependence thus produced. Since the relation of necessary connection across states is the hallmark of belonging to a common subject, whatever faculty or faculties are involved can be called 'transcendental (necessary for cognition) apperception' (self-consciousness, because the states consciously combined are states of the subject). What the faculty or faculties produces can be called the transcendental (necessary for cognition) unity of apperception (the relation of recognized necessary connection across the states of the subject that have been combined). These are not hypotheses about particular psychological faculties, but merely informative labels that point to the significance of functionally characterized faculties for a central philosophical problem.

Although Kant's investigation of the philosophical implications of conceptual cognition does not posit psychological capacities, it depends on and bears on cognitive psychology. Future research could undermine his philosophical results by showing that 'conceptual' cognition is not as it seems. Perhaps people (as well as oxen) do not understand grounds of cognition as such. Perhaps if asked they would deny that the concept application depends on other representations. Some two-year olds can 'count' in the sense that, if asked, they can correctly give the cardinality of a set of objects: 1, 2,

3, 4, 5. Other two-year olds can separate the ‘answer’ (the judgment) from the procedure; they repeat the value when they finish: 1,2,3,4,5 ...5. Even these more sophisticated counters don’t seem to see a relation between the procedure and the answer. When asked to give the number for the set of objects that they have just counted, they repeat the procedure: 1,2,3,4,5 ...5 and will do so again and again. By contrast, when asked about a set that they have already counted, 3-year-olds just give the answer: 5 (Karmiloff-Smith, 1992, 102). Presumably they understand at some level that they have already carried out the procedure that permits them to answer. If all ‘conceptual’ cognition were like the counting of 2-year olds—if it involved no recognition of the relation of some mental states to others—then Kant’s appeal to it to solve the philosophical problem of the unity of the ‘I-think’ would collapse.

Kant’s results can offer guidance to cognitive psychology. He gives partial accounts of what is involved in conceptual cognition and in having a unified self-consciousness and then establishes a connection between the two capacities. Burge accepts David Marr’s schema for the construction of complete accounts of psychological capacities, which includes a high-level description of a capacity (2010, xvii-xviii). So he might not object to using Kant’s analyses to give direction and constraints for fuller accounts of important psychological capacities.

I have tried to rebut the obvious objections that Kant’s theory of the transcendental unity of apperception credits laymen who are capable of cognizing with a philosophical theory of cognition and that it does so by invoking a conscious—yet completely unsuspected—faculty. These defenses are not intended as final and we’ll return to the issues in weighing Kantian criticisms of other approaches. We’ll also see some shortcomings of his views—although these do not, I argue, undermine the critiques.

### *3. Kant’s Theory and Anscombe’s Experiment*

In ‘The First Person’, Anscombe argues that ‘I’ cannot refer, because if it did, then the only possible referent would be a Cartesian ego, i.e., the thinking that thinks this thought (1975, 58). But such a referent cannot have temporal duration. For example, in a sensory deprivation tank, there is no way to identify the thinking that thinks this thought at  $t_1$  with the thinking that thinks this thought at  $t_2$  and so forth. If Kant is right

that cognizing through concepts requires making and recognizing necessary connections across different states, then Anscombe's thought-experiment is misdescribed. Since her thought, 'I won't let this happen again,' is a case of conceptual cognition, she must tacitly understand its dependence on other states, specifically, on the very odd state of having no current sensory stimulation, on her recollections of previous sensory states and so forth. She could not be in the situation where she refers to the thinking that thinks only this (current) thought, 'I won't let this happen again.'

On one point, Kant agrees with Anscombe. He argues that it is impossible to identify the same Cartesian ego, understood as 'the thinking that thinks this thought,' across different thoughts. Such 'thinkings' would be come and go and have no reference to a common subject. This was just his criticism of (probably) Locke in the B Deduction. He would object, however, that Anscombe presents a false picture in suggesting that there can be 'thinking' and 'I's' whatever the circumstances. In his view, thinkers are not born, but 'self-created' through exercising their powers of combination on materials that are combinable by them in judgments—through bringing different representations to the unity of apperception. That is why it is impossible, in principle, to present a legitimate use of 'I' while still leaving open the question of the unity of the 'I' or 'I-think.' If there is a thinker—if conceptual cognizing is going on—then different representations belong, and must be recognized as belonging, to a common thinker. Cognizers identify the I-think across different representations because they combine representations, because they think.

A defender of Anscombe could argue that she offers the characterization 'the thinking that thinks this thought' to acknowledge the fact that the use of 'I' cannot be mistaken. It cannot fail to refer and it cannot refer incorrectly (1975, 56-57). Kant has not explained how 'I's infallibly refer to themselves. Even if a conceptual cognizer must recognize different states as necessarily connected and so as belonging to a common subject, how, on Kant's view, is she justified in attaching 'I-think' to the representations?

Kant has an explicit and incorrect answer to this question: The thinker is aware of the states through inner sense (A22/B37). Since he relies on inner sense to distinguish a subject's mental life from the rest of the world, he does not address the question of a criterionless and so infallible referring use of 'I.' It is not clear whether this part of his

theory is consistent with his claim that cognizers cannot be aware of their identity through inner sense (A107). Against the inner sense theory, he argues that cognizers can recognize their identity through time only through conscious acts of combining that enable them to recognize different states as necessarily connected. Anyway, the inner sense theory is a bad theory, so this part of his view needs to be replaced. That is why I call the critique 'Kantian' and not 'Kant's.'

Although Kant would explain self-ascription through inner sense, his account of higher cognition offers a different way to approach the problem. He emphasizes that judging and inferring require conscious acts. Judgments and conclusions could be self-ascribed through those acts. We can express this point in terms of an entitlement principle:

When a cognizer (consciously) performs an act of higher cognition A that produces a state B, then she is entitled to self-ascribe A, and so B.

The consciousness involved is different from that of inner sense, because it is inseparable from acts of judging and inferring. A subject who consciously combines some representations in another that she recognizes as dependent on those representations cannot be mistaken that she is applying a concept or making an inference, not because the consciousness is a magical kind of perception, but because that is what, depending on the details, concept application is or what inferring is. Because the consciousness is part of what makes the act an act of higher cognition, the cognizer will always be in a position to self-ascribe the judgment or conclusion. It will always be possible for her to attach the 'I-think' to the judgment or conclusion and the self-ascription will always be correct—so she is entitled to self-ascribe. (I'll consider a skeptical objection in the last talk.)

I've argued against Anscombe's conclusion that 'I' cannot refer, because it is impossible to account for both the infallibility of the use of the term 'I' and the temporal duration of I's. My purpose was to reply to the objection that Kant's theory is too problematic to be used to criticize theories of self-knowledge. Although he was mistaken in relying on inner sense, this piece of the theory can be replaced with something suitable to the task. My target is not, however, Anscombe's conclusion that

'I' does not refer, but her often cited sensory deprivation tank thought experiment. If Kant is right, then the description of the case is inconsistent. She envisions a situation where she can self-ascribe a conceptual cognition without first understanding its relation to other representations in a single subject. Many discussions introduce the topics of self-knowledge or self-consciousness through the three features of infallibility, authority, and immediacy. In so doing, they assume that these issues can be resolved without first (or possibly ever) considering the unity of consciousness. Kant's theory of the unity of apperception implies that this assumption is mistaken. Further, on the 'Kantian' account of the infallible self-ascription of judgments and conclusions—that judging and inferring are a matter of consciously combining some states in others and of recognizing that combination as such—subjects who are capable of the infallible self-ascription of beliefs must also be aware of the unity of different states in a single subject.

---

<sup>i</sup> Here she makes the Fregean assumptions that the referent of a term is determined by its sense and that the sense is the way in which the object is presented to the person [1975, 55]

<sup>ii</sup> This is not so in cases of Lucy O'Brien and Chris Peacocke.