

## KANT, SELF-AWARENESS AND SELF-REFERENCE

Andrew Brook

### Introduction

As is well-known, Castañeda (1966, 1967), Shoemaker (1968), Perry (1979), Evans (1982) and others urge that awareness of self has peculiar features. It is less well-known that some of these peculiarities were discovered as early as 1781 and the *Critique of Pure Reason*.<sup>1</sup>

Two of the key peculiarities are that,

1. In certain kinds of awareness of self, first-person indexicals (I, me, my, mine) cannot be analysed out in favour of anything else, in particular anything description like,

and that,

2. In such cases, awareness of self is via what Shoemaker calls self-reference without identification. One can be aware of something as oneself without identifying it (or anything) as oneself via properties that one has ascribed to the thing.

(2) is often taken to be closely related to another putative peculiarity of awareness of self that Shoemaker (1970) calls

3. Immunity to error through misidentification, the idea that in some situations, we cannot become aware of a person by being aware of certain experiences, take that person to be oneself, and be wrong.

Shoemaker claims to have found the core of the idea in Wittgenstein (1933-4: 66-70).

There are questions to be asked about each of these three ideas. About (1): Is there not an ineliminable indexical element in most or all reference? About (3): Is there any such immunity and, if there is, how closely it is linked to (2) self-reference without identification? Certainly (3) has received a lot of attention in recent literature on self-reference and self-awareness; perhaps it has received more attention than it deserves. About (2): Isn't it the most basic of the three ideas? If (2) is the case, that would on the face of it seem to be enough by itself to produce (1) and (3). These are good questions and we will have something to say about all of them eventually. But

---

1. I explore the issues in this paper at greater length in Brook 1994: Ch's 4, 7 and 8. Unless otherwise noted, references to Kant are to the *Critique of Pure Reason*, in the Akademie pagination using the standard 'A' and 'B' notation for the first two editions (the only two that Kant prepared himself). A reference to one edition only means that the passage in question appeared only in that edition. Translations start from Norman Kemp Smith's 1927 translation and Guyer and Woods' 1998 translation and have been checked against the original text.

first we have a long, detailed piece of digging to do.

One standard argument for (1), that certain indexicals are essential, goes as follows.<sup>2</sup> To know that *I* wrote a certain book a few years ago, it is not enough to know that someone over six feet tall wrote that book, or that someone who teaches philosophy at a particular university wrote that book, or ... or ... or ... , for I could know all these things without knowing that it was *me* who has these properties (and I could know that it was *me* who wrote that book and not know that any of these things are properties of me). Nor would it help to add details of a more identifying kind — the person whose office number is 123 in building ABC, the person whose office phone number is ... . If I don't know that that office is my office, that that phone number is my phone number, I could know all these things and still not know that it was *me* who wrote the book. And vice-versa — through bizarre selective amnesia, I could *cease* to know all such things about myself and yet continue to know that it was *me* who wrote the book. As Shoemaker puts it,

... no matter how detailed a token-reflexive-free description of a person is, ... it cannot possibly entail that I am that person [1968: 560].

What a curious piece of knowledge – if it even is knowledge!

The standard argument for (2), that certain references to self do not require descriptive identification, goes as follows:

My use of the word 'I' as the subject of [statements such as 'I feel pain' or 'I see a canary'] is not due to my having identified as myself something [otherwise recognized] of which I know, or believe, or wish to say, that the predicate of my statement applies to it [1968: 558].

Whether Kant was aware of (1) is an intriguing question, as we will see. He was clearly aware of (2). Consider this passage:

In attaching 'I' to our thoughts, we designate the subject ... without noting in it any quality whatsoever — in fact, without knowing anything of it either directly or by inference [1781: A355].

This “attaching 'I' to our thoughts” business is interesting. In the kind of self-awareness in question, one is aware not just *of oneself*. One is aware of oneself *as oneself*. One is aware *that it is oneself* of which one is aware. Put in Fregean terms, one is not just aware of the being who happens to be oneself. One is *presented* to oneself in a certain way, namely, as oneself (see Ezcurdia, 2001). By “attaching 'I' to our thoughts”, Kant seems to have had something like this in mind, something like “using 'I' to refer to myself as the subject of my thoughts”. Since, on Kant's view, it is not just identifying properties but any properties whatsoever that I need not know in order to refer to myself as myself,<sup>3</sup> ‘non-ascriptive reference to self’ might capture what is special about this form of awareness of self better than Shoemaker's ‘self-reference without identification’ (Brook 1975: 188).

---

2. Though most often associated with Castañeda and Perry, a version of it can be found even earlier in Nagel (1965).

3. Being myself is not a property of me, i.e., something that I and other things could have in common.

But (2) is logically linked to (1). If I am aware of myself as myself without inferring this from anything else that I know about myself, my knowledge that it is myself of whom I am aware has to be independent, at least in some respects, of knowing anything else about myself. I can be aware of myself as myself without being aware of myself as anything except — myself. This shows that (2), of which Kant was clearly aware, requires something like (1). The existence of such a link between (1) and (2) does not establish that Kant was aware of (1), of course, but is still interesting. As we will see, there is some evidence that he was.

Moreover, Kant went further with these issues than any contemporary theorist. Unlike theorists of the past few decades, Kant had the makings of a theory to explain (1) and (2). Peacocke (2001) is the first contemporary theorist to have anything like such a theory.

Kant called the mode of reference to self that gives rise to awareness without identification *transcendental designation*. In Section 1 we will explore this notion to see what exactly it amounts to and how much Kant knew of the kind of reference that we use to achieve it. Section 2 lays out Kant's understanding of how awareness of oneself as oneself is different from awareness of one's psychological states (the latter is what Kant usually meant by 'inner sense') and oneself as an object among other objects. Section 3 lays out a further key element of Kant's theory, the notion of what I call the *global representation*. Kant held that one appears to oneself as a single common subject of a large number of psychological states (1781: A350). The notion of the global representation is closely related to that idea. Section 4 draws the elements of the theory together and in Section 5 we show how it explains (1) and (2). The theory turns out to have implications for some other things, too, immunity to error through misidentification in particular.

## 1. Transcendental Designation: The Referential Base of Self-Awareness

Kant's most common way of describing the awareness of self yielded by transcendental designation is this: "... through the 'I', as simple representation, nothing manifold is given" (B135). If the kind of reference that yields awareness of oneself as oneself, awareness of oneself as subject, is non-ascriptive, then the resulting awareness will not, or certainly need not, present any properties of oneself.

If awareness of self as subject is non-ascriptive, the reason, of course, is not that the self is some strange, indefinable being; as Kant brilliantly discerned, it is because of the nature of the acts of reference used to gain this awareness. Kant spoke of this kind of referring only a few times but when he did, he achieved insights into it that have only been rediscovered in the past thirty-five years.<sup>4</sup>

---

4. Transcendental designation, it is worth pointing out, is purely an epistemological phenomenon. Note, too, that even if this form of designation is transcendental, i.e., to do with the necessary conditions of experience, it is not transcendent, i.e., beyond experience. Whether Kant also held that something about the mind is transcendent is an ontological question and another matter altogether. Peacocke (2001) urges that Kant held the ontological thesis, too, but this is not clear. Did Kant think that, in addition to the mind of which we are aware in, e.g., acts of

For Kant's discoveries about reference to self, the crucial text is the passage on A355 of the first-edition attack on the second Paralogism quoted earlier:

In attaching 'I' to our thoughts, we designate the subject ... only transcendently, without noting in it any quality whatsoever — in fact, without knowing anything of it either directly or by inference.

This doctrine of transcendental designation also appears at B155 of the second-edition Deduction, though not thus labelled.

the I that I think is distinct from the I that it ... intuits ...; I am given to myself beyond that which is given in intuition [B155]

The central notion is the idea that I can refer to myself (presumably, as myself) using 'I' without 'noting ... any quality' in myself. One can refer to oneself in a variety of ways, of course: as the person in the mirror, as the person born on such and such a date in such and such a place, as the first person to do X, and so on, but one way of referring to oneself is special: it does not require identifying or indeed any ascription to oneself. So Kant tells us. Kant does not specifically say that transcendental designation makes me aware of myself *as myself* but he seems to have had it in mind. Indeed, as we will see a bit later, Kant argued that I must presuppose that it is myself of which I am aware to know some other things about myself.<sup>5</sup>

Did Kant really discover central features of self-reference without identification or did he merely happen onto something — something that we can identify in retrospect as self-reference without identification — with no more than a vague idea of what he had stumbled upon? Two things make me think that he had some definite idea of what he had uncovered.

First, he did not refer to the notion only once or in one way. He marked the basic claim in a number of ways and a number of places. He spoke, as we saw, of uses of 'I' that designate "only transcendently", "without knowing anything of [the subject]" (A355). But he also said that such uses "denote" but do not "represent" (A382). And, he tells us, even though the referent of 'I think' is as ubiquitous in experience as the referent of any categorical concept, it is not like other concepts because "it can have no special designation" (A341=B399). That is to say, presumably, it does not pick out its referent as one kind of object rather than another or as one object rather than another at all. (This last idea will be central in the fourth section.)<sup>6</sup>

---

transcendental designation, there is another, transcendent mind beyond such awareness? There is little reason to think that he did (Brook 1994, Ch. 4:5). To be sure, he thought that there is a lot about the mind that we cannot know — but here he meant the ordinary mind of self-awareness and introspection, not some second, transcendent entity.

5. One can compare what Kant says about reference to self to his doctrine that existence is not a predicate (A598=B626). In the same way that being aware of something's existence is not to be aware of any quality of it, being aware of oneself as oneself is something over and above being aware of qualities of oneself. In his criticism of Leibniz's Amphiboly, Kant says much the same thing about space and time — to be aware of space and time is to be aware of something over and above the qualities of space and time (A276=B332; see A281=B337).

6. Note that all the phrases just cited are from no earlier than the Paralogism chapter of the first edition. Kant seems not to have developed his theory of reference to self until he needed it to attack rational psychology. Note, too, that they are all from the first edition. In the second

Second, what if Kant's work also has the basics of the idea that certain indexicals are essential? That would be a good indication that he understood at least some aspects of reference to self without identification fairly well, because the latter phenomenon requires the former: indexical reference to self could not be essential unless there is a way of doing such acts of reference that is independent of (non-indexical) identification.

Whether Kant was aware that, for certain purposes, use of indexicals like 'I' is essential is not easy to settle with perfect confidence. Such evidence as there is, however, points to the conclusion that he was. Specifically, Kant argues that awareness of certain things presupposes awareness of oneself as "subject of the categories" in a way quite reminiscent of Shoemaker's claim that:

no matter how detailed a token-reflexive-free description of a person is, ... it cannot possibly entail that I am that person [1968: 560].

Here is what Kant says:

The subject of the categories cannot by thinking the categories [i.e. applying them to objects] acquire a concept of itself as an object of the categories. For in order to think them, its pure self-consciousness, which is what was to be explained, must itself be presupposed [B422].

This passage is from the extremely obscure second edition version of the Paralogisms chapter but the phrase 'its pure self-consciousness' seems to refer to awareness of oneself as oneself, awareness of oneself as subject. If so, what it seems to be saying is something like this: Judgments about oneself, i.e., ascriptions of properties to oneself, 'presuppose ... pure self-consciousness', i.e., awareness of oneself via an act of ascription-free transcendental designation. We find what may be the same claim in the first edition: "it is ... very evident that I cannot know as an object that which I must presuppose to know any object ... ." (A402).

A second passage, this time from Kant's introduction to the chapter on the Paralogisms, offers similar suggestions. The passage begins with a famous variant on the 'no manifold' theme that we saw at the beginning of this section:

Through this I or he or it (the thing) which thinks, nothing further is represented than a transcendental subject of the thoughts = X.

and then goes on:

It is known only through the thoughts which are its predicates, and of it, apart from them, we cannot have any concept whatsoever, but can only revolve in a perpetual circle, since any judgment upon it has always already made use of its representation [A346=B404].

The last clause is what interests us: "any judgment upon it has always already made use of its representation". Kant seems to be saying two things. On the one hand, to know that anything is

---

edition, Kant moved his discussion of awareness of oneself as subject to the Transcendental Deduction — and, regrettably, deleted most of the interesting details. A different but equally stripped-down version can be found in "The Psychological Idea," §46 of the *Prolegomena* (1783; Ak. IV:333-4).

true of me, I must first know that it is me of whom it is true. That would seem to require some awareness of myself that could not be derived from any knowledge of (what are in fact my) properties. Furthermore (a point that will become vital later), this awareness of myself as myself is not via some independent representation of myself. The only representations involved are representations of the objects of which I am aware. Awareness of self is somehow part of representations of objects. We do not have '*Ich-Vorstellungen*' of the sort envisioned by Frege or Husserl.

In short, these two passages seem to be saying something like this. In order to apply the categories to oneself, i.e., in order to make 'any judgment upon' oneself or know oneself as an object, one must already and independently be aware of oneself as subject, i.e., as oneself. But this is nothing less than the core of the idea of the essential indexical.

To summarize our results so far. Kant seems to have been aware of two features of reference to self that Shoemaker views as distinctive:

- (1) Kant was clearly aware of what Shoemaker calls reference to self without identification; in his jargon, we designate the subject "transcendentally, without noting in it any properties whatsoever" (A355); and,
- (2) There are indications that Kant was also aware of the idea of the essential indexical. In his terms, awareness of properties as properties of oneself presupposes awareness of oneself as subject, as oneself.

If so, there is reason to think that Kant did know what he had found when he hit on transcendental designation.

Few of Kant's students have paid much attention to transcendental designation. Doubtless there are many reasons for this. Kant managed to give a *clear* statement of what he thought about the topic only once in the whole *Critique*, namely, on A355. Not only is this description exceedingly brief but he dropped both it and the term in the second edition. It is also buried in the middle of an obscure and what many, I think wrongly, take to be a parochial discussion of the simplicity of the soul. Perhaps most importantly, the remarkable insights into the mechanics of reference to and awareness of self that Kant sketched there were lost again with his death, to reappear at the earliest with Wittgenstein (1933-4: 66-70) in his notion of the use of 'I' as subject and probably not until Castañeda and Shoemaker.<sup>7</sup> Indeed, it would have been difficult for anyone to have recognized what Kant had spotted prior to Castañeda's and Shoemaker's work. Kant himself lacked the apparatus needed to describe his discoveries adequately.

## 2. The Sources of Self-Awareness.

---

7. Though Shoemaker attributes the core of his treatment to Wittgenstein, one wonders. In his later writings Wittgenstein seems to have maintained that apparently self-referential uses of 'I' and cognates in fact are not referential at all. It is hard to tell whether he held the same view in his middle period. Perhaps we could put it this way: Kant anticipated what may be one thread in Wittgenstein's middle period work; Shoemaker developed the idea that Kant had anticipated, probably without knowing that Kant had worked on it.

Kant was not merely aware of some of the distinctive features of non-ascriptive self-awareness. Unlike recent theorists prior to Peacocke, he had the makings of a theory to explain them. Kant held that one gains awareness of oneself as oneself, as subject, in a way very different from the way in which one gains awareness not just of external objects but also of one's own psychological states (and even of oneself when one is an object of one's own mental states — for example, when one sees oneself in a mirror). He also held that the awareness of self that results is different from awareness of anything else in certain respects, respects that explain the resulting peculiarities. Let us start with how one becomes aware of oneself and what the resulting awareness is like.

Kant's theory of self-awareness compares and contrasts in interesting ways with some contemporary views. We just suggested a parallel between Kant's views and Peacocke's and we will return to it. Kant's views also contrast with some current theories, Rosenthal's higher-order thought theory and Dretske's displaced perception theory in particular. These contrasts bring out some of the originality and power of Kant's theory, as we will see when we get to them..

Since the distinctions just canvassed are part of Kant's doctrine of inner sense, we should start with an account of his doctrine of inner sense but we won't. His doctrine of inner sense is a mess. Here are just a few of the problems. Kant insists that all representational states are in inner sense, including those representing the objects of outer sense (i.e., spatially located objects),<sup>8</sup> but he also says that the object of inner sense is the soul, the object of outer sense the body (including one's own). He comes close to denying that we can be aware of the denizens of inner sense — they do not represent inner objects and have no manifold of their own. Yet he also says that we can be aware of them — representations can themselves be objects of representations — and that representations can make us aware of themselves. In its role as a form of or means to awareness of self, apperception ought to be part of inner sense. Yet Kant regularly contrasted apperception, a means to awareness of oneself and one's acts of thinking, with inner sense as a means to awareness of — what? Presumably, particular representations: perceptions, imaginings, memories, etc. Here is a passage from the *Anthropology*:

§24. Inner sense is not pure apperception, consciousness of what we are doing; for this belongs to the power of thinking. It is, rather, consciousness of what we undergo as we are affected by the play of our own thoughts. This consciousness rests on inner intuition, and so on the relation of ideas (as they are either simultaneous or successive). [1798, Ak. VII:161]

And these are just the most obvious problems. As I said, Kant's doctrine of inner sense is a mess.

Nevertheless, the quotation just introduced contains three interesting ideas. The first starts from the distinction between awareness of oneself and awareness in 'inner intuition' of 'what we undergo', i.e., awareness of our representational states. One of the things that we undergo in inner intuition is representation of oneself. If so, there is an important distinction between

---

8. "Whatever the origins of our representations, whether they are due to the influence of outer things, or are produced through inner causes, whether they arise *a priori*, or being appearances have an empirical origin, they must all, as modifications of the mind, belong to inner sense." (A98-9).

“apperceptive” awareness of self and awareness of self in intuitions. As Kant puts it in the second-edition Deduction,

... the I that I think is distinct from the I that it, itself, intuits ...; I am given to myself beyond that which is given in intuition, and yet know myself, like other phenomena, only as I appear to myself, not as I am ... [B155].<sup>9</sup>

This distinction between the ‘I that I think’ and the I that this I ‘intuits’ is the distinction between being aware of oneself as subject of all of one’s representational states and being aware of oneself as the object of some of them.

The second idea is the suggestion that apperceptive awareness of self is “consciousness of what we are doing” — *doing*. For Kant, we become aware of objects of representation via apperceptive acts of synthesis tying a manifold of intuition into coherent, recognizable, reidentifiable particulars. However, the standard way of becoming aware of an *act of representing* is quite different from this. We become aware of acts of representing not by receiving intuitions but by doing them: “... synthesis ..., as an act, ... is conscious to itself, even without sensibility” (B153); “... this representation is an act of *spontaneity*, that is, it cannot be regarded as belonging to sensibility” (B132). Kant tells us that here we do not represent by forming an object. That is not to say that acts of representing are not themselves represented in inner sense; they certainly are. It is just that they are not the *object* of a representation.

We can be aware of acts of representing via intuition, too, of course, but what is special about them is that we can be aware of them just by doing them, just by representing something. Another Kantian way of capturing what he had in mind here would be to distinguish between ‘awareness by doing’ and ‘awareness by having an image’ (for Kant, we are aware of all intentional objects in images [A120] — he even thought that all intentional objects are represented spatially [B154-5, B156, B158-9]). The point he is making is that we do not need to represent acts of representing in images. The passage from B153 cited above makes that clear.

Equally, we can be aware of ourselves as subject just by doing acts of representing. When I am aware of myself as the subject of a representation, I am aware of myself as doing the act of representing. One can of course be aware of oneself via intuition, too — by seeing oneself, for example, in a mirror. This is the way in which one becomes aware of one’s size, shape, colour, etc. But when one is aware of oneself as the subject of one’s own representations or agent of one’s acts, it is by being aware of acts of representing by doing them. Here is a passage from later in the *Critique* where Kant says this very clearly:

Man, ... who knows the rest of nature solely through the senses, knows himself also through pure apperception; and this, indeed, in acts and inner determinations which he cannot regard as impressions of the senses [A546=B574].

How does awareness of our own acts of representing and of ourselves as their subject work? The act of representing makes us aware of three things. Consider the sentence:

1. I am looking at the words on the screen in front of me.

---

9. In German, the first part is, “*Wie aber das Ich, der Ich denke, von dem Ich, das sich selbst anschaut, unterschieden ...* .” Kemp Smith translates this, “How the ‘I’ that thinks can be distinct from the ‘I’ that intuits itself ...” and inserts some unhelpful emendations.

Kant's claim seems to be that the representation of the words on the screen is all the experience I need to be aware not just of the words and the screen but also of the act of seeing them and of who is seeing them, namely, me. A single representation can do all three jobs. In Kant's words, the awareness of the latter two items is given "not indeed in, but with ... intuitions" (B161).

Let us introduce a term for this function. Let us call an act of representing that can make me aware of its object, itself and myself the *representational base* of my becoming aware of these items. Almost any representation will do. Imagining Pegasus will do just as well as perceiving external objects such as computer screens. Indeed, representational states which have no apparent object such as pains or feelings of hunger will do just as well. Nor does a representation itself have to be recognized to provide a representational base for self-awareness. Just recognizing the *object* of a representation is enough for me to be aware that it is me who is aware of it. Having the representational base for recognition of a state is not actually recognizing it, nor indeed myself, but it is to have all the *representation* I need. This is why, to return to a point made earlier, the basis of awareness of oneself is not some separate *Ich-Vorstellung*.

This theory of self-awareness is remarkably powerful. It rests on an idea that next saw the light of day in any clear form in Peacocke's (2001) theory. In the delta theory, having a property is sufficient in relevant cases for its subject to refer to and ascribe the property to itself — just Kant's most basic idea. Kant's theory also neatly avoids some of the problems that afflict other leading current theories of self-awareness.

Consider, for example, Rosenthal's (1991) higher-order thought theory. According to this theory, to be conscious of some representational state, A, one must have another representation (a thought) of which A is the object. For Kant, all we need is A itself. A representation itself has the power to make us aware of it. Rosenthal's view runs foul of objections such as that lots of creatures conscious of their own states don't seem to have such complicated thoughts, and that the model readily leads to a regress of thoughts about thoughts about thoughts, and there is little independent reason to postulate such a hierarchy of thoughts. Now of course Rosenthal has answers to such objects — but for what we might call Kant's *same-order model*, no such objections arise. Advantage Kant.

Or consider Dretske's (1995) displaced perception view of self-awareness. Here when we are aware of our own representations (and, by extension, ourselves as their subject?), this is a matter of inference from what we are representing, that this thing is being represented. Hence his term 'displaced perception'. This theory runs foul of the objection that we seem to know a lot about our representations that goes well beyond what is being represented in them, e.g., whether they are striking or faint, in one sensible modality or another, perception, memory or imagination, and so on and so forth. Despite a superficial resemblance engendered by the idea that representation of an object is all we need for awareness of the representation itself, Kant's theory is not a displaced perception theory. Indeed, nothing is displaced from the object being represented. Representation of the act of representing and representation of self are distinct activities. It is just that a single representations provide for all three; we do not need a distinct representations for each act. Once again, Kant is safe from objections that face others. Once again, advantage Kant.

Now the third idea introduced in the passage quoted earlier. It is this. When one is aware of oneself as subject by doing acts of representing, one is not aware of oneself as an object of

representation of any kind. (By ‘object’ I mean intentional object.) Here are some passages: “it is ... very evident that I cannot know as an object that which I must presuppose to know any object ... .” (A402). “[The representation] ‘I’ is ... as little an intuition as it is a concept of any object” (A382). “The proposition, ‘I think’, in so far as it amounts to the assertion, ‘*I exist thinking*’ ... determines the subject (which is *then* at the same time object) in respect of existence” (B429; second emphasis mine). To be aware of myself as an object requires not only “spontaneity of thought”, that is, acts of transcendental apperception, but also “receptivity of intuition”; that is, it requires “the thought of myself applied to the empirical intuition of myself” (B430-31).

Let us spell this idea that awareness of oneself as subject as not awareness of oneself as an object of a representation out a bit. Kant tended to tie awareness of objects very closely to sensibility, to appearances and intuitions (see A104); all awareness of objects seems to be via sensibility. But, as Kant put it, “... synthesis ..., as an act, ... is conscious to itself, even without sensibility” (B153); being aware of an act of representing by doing it is not being aware of it by receiving intuitions of it. In addition, he says that the “unity of the synthesis of the manifold” (i.e., that we are representing the manifold in a single representation) is given “not indeed in, but with ... intuitions”. Since one is aware of oneself as subject by being aware of acts of representing by doing them, not via intuitions in sensibility (B430), it would follow that one is not aware of oneself as a represented object. (On the other hand, Kant does say that the subject is an object, the ‘transcendental’ object of inner sense, at A341 and A361. Because he uses *Gegenstand*, not *Objekt* on those occasions, I take him to be talking about objects in a loose way here that does not negate the distinction so carefully drawn in the passages quoted above.)

For Kant, this distinction between awareness of self by doing acts of synthesis and awareness of things as objects is of fundamental importance. When one is aware of oneself by doing cognitive and perceptual acts, one is aware of oneself as spontaneous, rational, self-legislating, free — as the doer of deeds, not just as a passive receptacle for representations: “I exist as an intelligence which is conscious solely of its power of combination” (B158-159), of “the activity of the self” (B68) (Sellars, 1970-1; Pippin, 1987).

To close this section, let me situate Kant vis-à-vis one more strand in recent thought about self-awareness, the so-called noncognitivist approach associated with the later Wittgenstein. Strangely enough, some recent commentators, Powell (1990 for example, have taken Kant’s assertion that awareness of self as subject is not intuitional awareness of objects to be an assertion that what seems to be reference to self as subject is not a referential act at all and even that there is nothing there to refer to. The idea is this. If all (noninferential) awareness of things is intuition of objects, then, if uses of ‘I’ are nonintuitional, they are either nonreferential or do not give us awareness of anything or both. I know of nothing in Kant that could support such an interpretation and it is not plausible in its own right. As Shoemaker says, “in making a judgment like ‘I feel pain’ one is aware of [no]thing less than the fact that one does, oneself, feel pain.” (1968: 563). I think these exotic readings of Kant rest on two mistakes. One is the mistake of thinking that, for Kant, all noninferential awareness is intuitional. The other is the mistake of confusing non-ascriptive reference with absence of reference.

To summarize this section: According to Kant, we have two ways of becoming aware of ourselves — by working intuitions up into intentional objects and by being aware of acts of representing and of oneself as their subject by doing them. The former makes us aware of

represented objects, the latter makes us aware of ourselves as the subject and agent of the act of representing. This distinction anchors Kant's theory of the peculiarities of awareness of oneself as subject. Before we can lay out that theory, we need to add one further element.

### 3. The Global Representation

So far we have focussed on individual representations. For Kant, however, the representations that serve as the representational base of awareness of oneself as subject are usually much 'bigger' than that, i.e., contain multiple objects and often multiple representations of them tied together into what Kant called 'general experience'.

When we speak of different experiences, we can refer only to the various perceptions, all of which, as such, belong to one and the same general experience. This thoroughgoing synthetic unity of perceptions is indeed the form of experience; it is nothing else than the synthetic unity of appearances in accordance with concepts [A110].

Call this general experience a *global representation*. When I am aware of many objects and/or representations of them as the single object of a global representation, the latter representation is all the representation I need to be aware not just of the global object but also of the global representation itself (the 'one consciousness' and the various individual representations that I have of objects), and of myself as not just the subject of individual representations but as the common subject of all the constituent representations.

One can be aware of more than one act and object of representing. However, to become aware of any such acts and/or objects is to integrate them with the acts and objects of representation of which one is already aware. At any one time, there will be one largest act of representing. That is one's current global representation. In short, the global representation is the home of the unity of consciousness, a unity that has both a synchronic and a diachronic dimension.<sup>10</sup>

According to the last paragraph, each of us could have only one global representation at one time. If so, to use a phrase of Wittgenstein's (1935-6), a global representation has no neighbour; there will not be other, simultaneous global representations of which one is aware by having them from which to distinguish any one of them. This uniqueness will become important when we draw out Kant's theory of the peculiarities of awareness of self in Section 4.

How do global representations serve as the representational base of awareness of self as subject? Here, stated with less than pellucid clarity, is Kant's account:

the mind could never think its identity in the manifoldness of its representations... if it did not have before its eyes the identity of its act, whereby it subordinates all [the

---

10. It is notoriously difficult to elucidate the kind of unity involved (Brook 1994, Ch. II: 5). However, noticing that it has both synchronic and diachronic dimensions can help us sort out where immunity to error through misidentification could occur and where it could not, as we will see later.

manifold] ... to a transcendental unity... [A108].

Kant seems to mean that what allows me to become aware of my identity as the common subject of my various representations is that I can be aware of the single, unified and unifying acts of representing by which I combine the objects of these representations, and sometimes also the representations themselves, into a single object of a global representation. Kant called these acts *transcendental apperception*. I think he was expressing the same idea when he said in the second edition that I am aware of myself as the single common subject of a certain group of experiences by being aware of “the identity of the consciousness in ... conjoined ... representations” (B133).

Put differently, Kant thought that synthesis into global objects is a necessary condition of awareness of self as common subject. Without objects of representation being tied together as a single complex object of a single representation, one might be aware of the subject of an individual representation but one could not be aware of the subject of one such representation as the subject of other such representations. Rather, I should have “as many-coloured and diverse a self as I have representations of which I am conscious ...” (B134) — as are *in fact* had by me, for I would not, of course, be aware that it was me. It takes a global representation to serve as the representational base of awareness of self as common subject.<sup>11</sup>

Is a global act of unifying representation also *sufficient* for being aware of oneself as subject? No; despite many claims to the contrary, notably by Strawson (1966), Kant was clear that one could represent objects without being aware of oneself (A113; A117fn.; B132). If one can have representations of which one is not aware, as was suggested earlier, one could have *global* representations of which one is not aware — for example, if one’s attention was totally focussed on the complex scene being represented. Even if each global representation is the full *representational base* of self-awareness, a direction of attention or some cognitive apparatus necessary for taking advantage of the available representational opportunities might be missing.<sup>12</sup> (Something like this might explain why nonhuman animals are not aware of themselves as subjects.)

To summarize. One’s global representation is the representational base of being aware of objects, of the representation itself, and of oneself as the common subject of one’s representations. We turn now to the theory of the peculiarities of self-awareness that can be built on this base. All the pieces of it are to be found in Kant.

#### 4. Why Awareness of Self is as It is

The basic idea behind the theory whose pieces can be found in Kant is fairly simple. When the medium of awareness of oneself is the doing of acts of representing, the medium imposes sharp constraints on what the resulting awareness can be like. These constraints account for the peculiarities of awareness of self as subject identified at the beginning of this paper. That Kant spotted distinctive features of “reference to self without identification” is remarkable enough. That he had the makings of a theory to explain some of them is even more remarkable.

---

11. Discussions with Richard DeVidi saved me from some errors here.

12. Kant was aware of attention but discussed it only once in the first *Critique* (B156 fn.).

As we saw in Section 2, the special awareness that we have of ourselves as subject is not the only form of awareness of self that we have but it is different from intuitional awareness of oneself as an object and the differences can explain some of the peculiarities of self-awareness.

This theory is neither easy to spot nor simple to unravel. Kant never laid it out completely, indeed he only hints at some of its most important features. Thus any reconstruction of it is bound to be speculative. The key component of the theory, not surprisingly, is the global representation. If A355 is the crucial text for Kant's view of awareness of self as subject and the kind of reference that yields it, A108 is the crucial text for the theory of why those things are as they are. Between them these two pages contain the core of Kant's whole picture of self-awareness. Regrettably, they could not be more obscure.

One can be aware of oneself by seeing oneself in a mirror, by acts and states of inner sense such as feelings, thoughts, etc., being objects of representations, by inferring things about oneself from other things of which one is aware, by being aware of oneself through perceptions of one's body, behaviour, etc. in outer sense (A347), and perhaps in other ways. In all these cases, one would be aware of oneself as the object of a representation.

To be aware of myself as an object, I need a representation devoted to that (intentional) object, a representation of which the thing that is me rather than something else is the object. Now return to the earlier claim that any representation that I am having can be the representational base of awareness of self as subject, no matter what its object is.<sup>13</sup> If that is right, then all representations of which I am aware by having them present the same subject to me, namely, me, and in the same way, as me. *A fortiori*, the single common subject of a global representation is presented to itself as the same subject throughout this global representation. If so, when one is aware of oneself as oneself, as subject, this awareness is not experience-dividing, to use a term of Bennett's — "i.e., [statements expressing it have] no direct implications of the form 'I shall experience C rather than D'" (Bennett, 1974: 80). In a statement such as

1. I am looking at the words on the screen in front of me,

the verb expression or the object expression may divide experience but the subject expression does not. In this, awareness of self as subject is unlike all other awareness.

Why is this kind of awareness of self not experience-dividing? As Kant puts it, "if I want to observe the mere 'I' in the change of all representations, I have no other *correlatum* to use in my comparisons except again myself" (A366; cf. A346=B404, B422). That is to say, I could not compare this 'I' to or contrast it with what is presented in any representation I am having in which it does not appear. To use Wittgenstein's phrase once more, awareness of oneself as subject has no neighbour. In no representation of which I am aware by having it does the subject appear differently from how it appears in any other.<sup>14</sup> If so, I could not distinguish the self

---

13. Even some representations that might be somebody else's by some other criterion seem to provide that possibility. This will complicate the business of immunity to error through misidentification when we get to it.

14. An obscure remark in the attack on the third Paralogism may be based on the same idea: "in the apperception, time is represented... only *in me*" (A362). Part of what Kant may have meant here is that all my representations that locate something in time or are themselves located in time will also represent myself.

presented in one such representation from the self presented in any other such representation.

We can now explain why, when one appears to oneself as oneself, one is not appearing as the object of a representation. A representation is individuated, differentiated from other representations, by its object. But no representation of mine is made different from any other representation of mine by the fact that it makes me aware of myself as its subject. To represent something as an object is to place it vis-à-vis other objects, and usually to ascribe properties to it. If so, to appear in a representation as subject is not to appear as an object of any kind, just as Kant said (A342=B400). ( Interesting enough, Shoemaker says something similar [1968: 563-4].)

The basis in theory for Kant's insistence that one's awareness of oneself as subject is not via "noting qualities" of oneself can be put this way. We can distinguish the subject from all objects (A342=B400). What we cannot do is compare it to, contrast it with, one object *rather than another*. If so, awareness of self as subject does not distinguish me from or identify me with anything of which I am aware as an object, anything in "the world". Something of great interest follows: so far as anything my awareness of myself as subject could tell me, I could be *any object or any compilation of objects or any succession of objects* whatsoever. Not by accident are these exactly the topics of the first three Paralogisms. One of the mistakes of rational psychology is precisely to take the simplicity (lack of manifoldness) in the unified representation of self to be a representation of simplicity and unity. Kant, of course, insists that awareness of self as subject tells us *nothing* about what the self is like (A355; B156).

## 5. Putting the Theory to work

Let us now put the idea of the representational base and the observations we have just made about awareness of self as subject to work on some of the peculiarities of awareness of self as subject discussed at the beginning of the paper. First, (2), self-reference without identification.

Two questions. First, how is self-reference with identification possible? It is possible because awareness of oneself as the subject of representations of which one is aware by having them (what a phrase!), one always appears to oneself in the same way. Thus, to recognize oneself as subject does not differentiate the entity thus recognized from anything else presented as subject. If a representation does not thus differentiate, does not divide experience, then one does not need to ascribe properties to it in order to achieve reference, indeed unique reference, to it. If representation can be without ascription in this way, then it is possible that, as Kant put it, "through the 'I', as simple representation, nothing manifold is given" (B135), that there is a way of referring to oneself that cannot be "accompanied by any further representation" (B132), in which 'I' can only 'denote', not 'represent' (A382), can designate "only transcendently" (A355), that can "have no special designation" (A341/2=B399). This awareness is a 'bare consciousness' (A346=B404; B158) that is not knowledge (B157).

Second question. Must self-reference take place without identification if one is aware of oneself as subject at all? It must. If I appear to myself in the same way in every representation of which I am aware by having it, not is awareness of self as subject without ascription *possible*, it *must* occur without ascription. Ascribing properties would produce just the differentiations that

are not there in awareness of self as subject. But this allows us to explain (2), self-reference without identification. If self-awareness requires reference to self, and if self-awareness takes place with no ascription, then reference to self must take place with no identification. If so, we have explained (2).

And also (1), the essential indexical. The idea of the essential indexical is that to make references to self via ascribing properties, one must be able to make references to self that do not ascribe properties. What alternative to ascriptive reference is there? So far as I know, only indexical reference. If so, indexical reference is essential to awareness of oneself as subject. Castañeda (1966) seems to hold that this feature of awareness of self as subject is simply a brute fact of self-awareness and cannot be further explained. Had Kant ever have thought about the issue, I conclude, he would have disagreed.

However, this does not explain everything that needs to be explained. Castañeda and Shoemaker argue not only that indexical references to oneself are essential, but that in them one must refer *to oneself as oneself* (or cognate). Does Kant have anything to explain this final feature of indexical reference to self? Not that I know of.

The materials that Kant left us can help with some other things, however. First, an element of his account shows that indexical reference using 'I' or cognates is essential in a narrower range of cases than Castañeda and Perry thought. They claim that from knowledge of properties by themselves, one could not know that the properties were one's own unless one already knew of oneself and *as oneself*. Notice that the way in which one is aware of the property is not specified here. If we become aware of certain properties in a certain way, the claim does not hold. Consider feeling a pain and becoming aware of it by feeling it. When I am aware of a pain by feeling it, I know or certainly can know that it is a pain *of mine*. Yet we described the pain non-token-reflexively ('the person involved is aware of the pain by feeling it'.) If so, the claim that indexical reference is essential holds in a narrower range of cases than has been thought.

Indexical reference is essential only in those situations in which *how* one is aware of a state or event does not settle *who has* that state or event. Roughly, how one is aware of a state or event settles ownership for psychological states where there are two or more ways of becoming aware of them, namely, by having them and by observing and inferring them in various ways and one is aware of the state by having it. The element in Kant's account that makes this narrowing possible is, of course, his distinction between awareness of a state by having it and awareness of it as the object of a representation.

This distinction, between awareness by having an experience and being aware in other ways, can also be put to work to explain — and again to limit — Shoemaker's notion of immunity to error. This, as we said, is the idea that certain judgments are "immune to error through misidentification with respect to the first person" (1970: 269-70; see 1968: 556) What Shoemaker meant, as we saw at the beginning of this essay, is that in some situations we cannot become aware of a person by being aware of certain experiences, take that person to be oneself, and be wrong. To take the most plausible kind of example, I could not be aware of seeing a scene, decide that it is me seeing it, and be wrong. Shoemaker introduced the notion in the context of past-tense memory judgments but if it ever obtains, it obtains most clearly in the present. (We will return to the past-tense cases in a moment.)

Kant's account suggests that if I am ever immune to error in identifying a person as myself, it could only be when I am aware of the person in question by *having* the experiences or *doing* the actions via which I am picking out the person. Clearly there is no such immunity when I am aware even of myself on any other basis, e.g., by looking in a mirror, seeing a body part, or hearing someone discussed.

Indeed, there may be no such immunity even in some cases where I *am* aware of experiences or actions by having or doing them. Suppose that one could be aware of a feeling by feeling it when this feeling, as judged by any other criterion, is someone else's. Feeling a feeling is one criterion for a feeling being one's own but it is not the only one. Perhaps changes in the feeling are causally dependent on another's body, the feeling has as its object a scene represented from the perspective of another person, another person (or at any rate another *body*) can report on the feeling in exactly the way that I can, and so on. If application of the various criteria for who is having that feeling produced a mixed result of this sort, then Shoemaker would not be obviously right to hold that we are immune to error through misidentification with respect to the first person even about some experiences and actions of which we are aware by having or doing them.<sup>15</sup> Nor does such a split seem impossible. One could imagine wiring that let one feel a feeling, perceive a scene, etc., where these feelings and perceptions were in every other respect associated with another person. They were seen through the other's eyes, the other person can report having and feeling them in exactly the same way and on the basis of exactly the same kind of experience as I can, and so forth. Simply being aware of a perception, feeling, etc., by having it does not render the question of whose perception or feeling it is trivial.

Still further, if immunity to error through misidentification with respect to the first person obtains at all, it obtains only of states, feelings, actions, etc., that one is currently undergoing. There is no immunity with respect to things that I remember having or doing. When I remember some thought or feeling or action from the point of view of thinking or feeling or doing it, I will "automatically assume" (Parfit 1970: 15) that it is me who thought or felt or did it; I will be in the same situation as if it were me that I am aware of. Yet, if traces of autobiographical memories (more exactly, autobiographical q-memories) can be transferred from person to person in certain ways, this need not be true. Suppose that I (q-)remember seeing the scene from the top of Mt. Robson. I will 'automatically assume' that it was me who saw that scene. Yet I have never been on top of Mt. Robson. Sally has, however, and memory-traces set up in her when *she* saw that scene have been transferred to me. If so, it is Sally that I am remembering seeing it, not me. I take it to be me because I remember the scene from the point of view from which she saw it but it is still her that I am (q-)remembering. If this analysis is right, there is no immunity to error through misidentification in autobiographical (q-)memory.<sup>16</sup>

The *ur*-phenomenon for Kant is the *manifoldlessness* of the self's appearance to itself.

---

15. Powell (1990) raises some doubts about whether such immunity is unique to first-personal reference, referring to Evans' similar thoughts (1982). He also challenges Anscombe's appeal to this immunity as a reason to think that uses of 'I', etc., are non-referential. Strangely, in the end he himself opts for a non-referentialist reading of Kant!

16. Evans (1982: 238) disagrees. He thinks that some autobiographical memories are immune to error through misidentification with respect to the first person. It seems to me that he runs two notions together, namely, awareness of self without identification and immunity to error through misidentification. They need to be kept separate.

Interestingly, this lack of qualitative manifold appears even phenomenologically. One can easily observe in oneself that having a representation gives one information about that representation (what type of representation it is, how it was formed, what we are representing by it, etc.) but it merely presents oneself as subject of it. True, the self is presented as me and this might be a mode of presentation of a sort (Ezcurdia, 2001) but it is at best an extremely stripped down mode of presentation. Beyond this ‘information’ that the thing presented is me, the representation tells me nothing about myself. This barrenness in one’s awareness of oneself as subject was perhaps one of the things that led Hume to think that no subject is to be found in self-awareness at all.

To summarize. Kant seems to have anticipated the idea of the essential indexical. He unquestionably anticipated the idea of reference to self without identification. And he sketched a most interesting theory of why it must be so. The theory is based on his claim that having a representation, any representation, is the representational base for awareness of self as subject. It seems unlikely that anything like immunity to error through misidentification would ever have occurred to him — but his theory can help us limit and think more clearly about this notion, too. Not a bad record for someone writing over 200 years ago!

Nor have we exhausted Kant’s contribution to our understanding of self-awareness, not by any means. For example, his distinction between awareness of self as subject via doing acts of apperceptive synthesis and awareness of one’s psychological states via representations in inner sense is very much like the distinction Evans (1982, Ch. 7) draws between nonidentificatory awareness of self and introspection. Again and also like Evans, Kant was interested in how awareness of oneself as subject and awareness of oneself as object among other objects connect. For example, how can we know that the thing of which we are aware as subject is one of the things of which we are aware as object (B155)? These issues deserves their own unhurried treatment. Here I hope to have shown that Kant anticipated some of the most important ideas on self-reference and self-awareness of the past thirty-five years.

## References

- Bennett, J. 1974. *Kant’s Dialectic*. Cambridge: Cambridge University Press
- Brook, A. 1975. Imagination, possibility and personal identity. *American Philosophical Quarterly* 12: 185-98.
- . 1994. *Kant and the Mind*. New York: Cambridge University Press
- Brook, A. and R. DeVidi. 2001. *Self-reference and Self-Awareness*. Amsterdam: John Benjamins
- Castañeda, H.-N. 1966. ‘He’: A study in the logic of self-consciousness. *Ratio* 8:130-57.
- . 1967. Indicators and quasi-indicators. *American Philosophical Quarterly* 4:85-100.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, MA: MIT Press/A Bradford Book.
- Evans, G. 1982. *The Varieties of Reference*. J. McDowell, ed. Oxford: Clarendon Press (Chapter VII)

- Ezcurdia, M. 2001. Thinking about myself. In Brook and DeVidi, 2001.
- Kant, I. 1798. *Anthropology from a Pragmatic Point of View* (Ak VII). Mary McGregor, trans.. The Hague: Martinus Nijhoff, 1974 ['Ak' stands for *Kant's Gesammelte Schriften*, ed. Koniglichen Preussischen Akademie der Wissenschaften, 29 vols. Berlin: Walter de Gruyter et al., 1902 - , commonly known as the Akademie edition].
- . 1781/87. *Critique of Pure Reason*. Norman Kemp Smith, trans. in *Immanuel Kant's Critique of Pure Reason*. London: Macmillan Co. Ltd., 1963. Also trans. Werner S. Pluhar. Indianapolis, IN: Hackett Publishing Co., 1996, and Paul Guyer and Allan Woods, Cambridge: Cambridge University Press, 1998. .
- Nagel, T. 1965. Physicalism. *Philosophical Review* 74: 339-56
- Peacocke, C. 1983. 2001. First-person Reference, Representational Independence, and Self-knowledge. In Brook and DeVidi, 2001. Also in Wolfgang Kunne, Albert Newen, and Martin Anduschus, eds. *Direct Reference, Indexicality, and Propositional Attitudes*. Stanford, CA: CSLI Publications.
- Perry, J. 1979. The problem of the essential indexical. *Noûs* 13:3-21. Reprinted in J. Perry. 1993. *The Problem of the Essential Indexical*. New York: Oxford University Press: 33-53.
- Pippin, R. 1987. Kant on the Spontaneity of Mind. *Canadian Journal of Philosophy* 17: 449-76
- Powell, C. T. 1990. *Kant's Theory of Self-Consciousness*. Oxford University Press.
- Rosenthal, D. 1991. *The Nature of Mind*. Oxford: Oxford University Press
- Sellars, W. 1970-1. " ... this I or he or it (the thing) which thinks ... ." *Proceedings of the American Philosophical Association* 44: 5-31
- Shoemaker, S. 1968. Self-reference and self-awareness. *Journal of Philosophy* 65:555-67.
- . 1970. Persons and their Pasts. *American Philosophical Quarterly* 7: 269-285.
- Wittgenstein, L. 1935-6. Wittgenstein's Notes for Lectures on 'Private Experience' and 'Sense Data'. *Philosophical Review* 77, 1968: 271-320.
- . 1933-4. *The Blue and Brown Books*, 2nd ed. Oxford: Blackwell, 1969.