

the review of  
**metaphysics**

a philosophical quarterly

---

Kant and Reductionism

Author(s): Quassim Cassam

Source: *The Review of Metaphysics*, Vol. 43, No. 1 (Sep., 1989), pp. 72-106

Published by: [Philosophy Education Society Inc.](#)

Stable URL: <http://www.jstor.org/stable/20128835>

Accessed: 01-02-2016 10:27 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*Philosophy Education Society Inc.* is collaborating with JSTOR to digitize, preserve and extend access to *The Review of Metaphysics*.

<http://www.jstor.org>

## KANT AND REDUCTIONISM

QUASSIM CASSAM

**I**N *REASONS AND PERSONS*, Derek Parfit defends a conception of the self or person which he labels “Reductionist.”<sup>1</sup> It is a conception which owes much to Hume’s view of the self as a bundle of causally connected perceptions.<sup>2</sup> Indeed, Parfit’s account might be thought of as capturing the best insights of the bundle theory, while avoiding many of the objections to which cruder versions of that theory appear to be liable. Parfit’s preliminary characterization of Reductionism is in connection with the notion of personal identity. A Reductionist view of personal identity holds that: (1) the fact of a person’s identity over time just consists in the holding of certain more particular facts, and (2) these facts can be described without either presupposing the identity of this person, or explicitly claiming that this person exists. These facts can be described in an *impersonal* way.<sup>3</sup>

A position is non-Reductionist if it rejects either of these claims. The main non-Reductionist account with which Parfit contrasts his own account is the Cartesian view that the person is a spiritual substance, a separately existing entity, distinct from his brain and body and his experiences.

In this paper, I will attempt to sketch and defend a view of the self which is neither Cartesian, nor, in Parfit’s sense, Reductionist, although it is in agreement with Reductionism on a number of important points. Parfit himself recognizes the possibility of such a middle position, and attributes it, with some justification, to Kant.<sup>4</sup> The Kantian response to Descartes’ conception of the self is not

---

<sup>1</sup> Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1986), pt. 3.

<sup>2</sup> See *A Treatise on Human Nature*, ed. L. A. Selby-Bigge, 2d ed. (Oxford: Oxford University Press, 1978), bk. 1, pt. 4, sect. 6.

<sup>3</sup> Parfit, *Reasons and Persons*, 210.

<sup>4</sup> *ibid.*, 225.

difficult to understand, but what Kant's view of the bundle theory was, or ought to have been, is far more difficult to assess. On one view, Kant's account of the unity of consciousness in the "Transcendental Deduction" is best understood as a declaration of an anti-Humean position on the self.<sup>5</sup> Hume himself conceded that his bundle theory was unable to provide for the unity of consciousness.<sup>6</sup> This is a suggestion which Kant would probably have viewed with some sympathy. But if Kant is to be viewed—as he surely must be—as attempting to steer between the Scylla of Cartesianism and the Charybdis of bundle theory, the scope of his disagreement with the Humean position needs to be clearly defined. If Kant has an argument against the bundle theory, it would be interesting to know what it is. In this paper, I will proceed as follows: I will begin by discussing Parfit's version of the bundle theory, in an attempt to identify clearly just what it is committed to, and the differences between it and cruder versions of the theory. I will then attempt to explain and defend Kant's claim that the unity of consciousness, as he understands it, is required if we are to have the capacity to conceptualize our perceptions as perceptions of mind-independent, spatial objects.<sup>7</sup> Finally, I will consider the question of whether the basic insight yielded by Kant's transcendental argument can be respected by the Reductionist, with his quite different conception of the unity of consciousness. If it turns out that even Parfit's highly sophisticated version of the bundle theory is incompatible with the Kantian account of how we are able to make sense of perceptions as perceptions of spatial objects, then, although this would not amount to a refutation of bundle theory, it would suggest that the theory is even more revisionary than might at first have been thought. Parfit's Reductionist takes for granted our conception of our experience as experience of an objective, spatial world, but for Kant, how we think of the self and how we conceptualize our experiences are intimately related issues. It remains to be seen whether or not there is any conflict between Reductionism and Kant's central "objectivity requires unity of consciousness" argument in the "Transcendental Deduction of the Categories."

---

<sup>5</sup> See Patricia Kitcher's "Kant on Self-Identity," *The Philosophical Review* 91 (1982): 41–72.

<sup>6</sup> See *A Treatise*, 636.

<sup>7</sup> See the "Transcendental Deduction" in the *Critique of Pure Reason* (London: Macmillan, 1929). In the deduction, Kant sets out to establish

## I

According to Parfit, any position which deserves to be called “Reductionist” must at least accept that (3) a person’s existence just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events. In addition, Parfit’s sophisticated Reductionist claims that (5) a person is an entity that is *distinct* from a brain and body, and such a series of events. Finally, a Reductionist must hold that (9) though persons exist, we could give a *complete* description of reality *without* claiming that persons exist.<sup>8</sup> Corresponding to (2) and (9) is the claim that it is possible to explain the unity of consciousness in impersonal terms. The unity of consciousness at a time and the unity of a whole life are both to be explained by describing the relations between experiences, and their relations to a given person’s brain: “We can refer to these experiences and fully describe the relations between them, without claiming that these experiences are had by a person.”<sup>9</sup> Whereas (5) is optional, (9) and the corresponding unity of consciousness appear, on Parfit’s view, to be constitutive of Reductionism.

How can (3) and (5) be reconciled? Parfit remarks that a person is not a composite object, but is rather an entity that *has* a brain and body, and *has* particular thoughts, desires and experiences. The sense in which (5) is true, and compatible with (3), is that we *speak* of persons as the bearers of mental states. Mental and physical states are ascribed to ourselves and others in the form of subject-predicate propositions, and it is for this reason that Parfit rejects the crude Reductionist thesis that (4) a person *just is* a particular brain and body, and a series of interrelated mental and physical events.<sup>10</sup> It is, as Wittgenstein might have put it, a feature of our “grammar” that I am not a bundle or series of events, and am in

---

a connection between the unity of consciousness and experience of independent objects. The connection between objectivity and spatiality is asserted in the “Transcendental Aesthetic.”

<sup>8</sup> Parfit, *Reasons and Persons*, 211–12.

<sup>9</sup> *ibid.*, 217.

<sup>10</sup> Given Parfit’s rejection of (4), he may dispute the characterization of his view as a bundle theory, although this characterization will be relatively harmless if its point is merely to reflect the Humean ancestry of Parfit’s theory. See note 18 below.

*this* sense distinct from a series of mental and physical events, but it does not follow from this that (3) is false.

Parfit's way of reconciling (3) and (5) also holds the key to his response to the claim that his form of Reductionist is what Strawson once called a "no-subject theorist." On one interpretation, the bundle theory is committed to the view that the belief in the existence of subjects of experience, to which mental states are properly ascribable, is, quite literally, an error. As Strawson wrote on behalf of the "no-subject" theory, "It is only a linguistic illusion that one ascribes states of consciousness at all, that there is any proper subject of these ascriptions, that states of consciousness belong to, or are states of anything at all."<sup>11</sup> In contrast, Parfit's Reductionist, who rejects (4) and accepts (5), denies that he is an error theorist, denies that there are not subjects of experience, to which states of consciousness are properly ascribed in the form of subject-predicate propositions:

Even Reductionists do not deny that people exist. And, on our concept of a person, people are not thoughts and acts. They are thinkers and agents. . . . A Reductionist can admit that . . . a person is *what has* experiences, or the *subject of experiences*. This is true because of the way in which we talk. What a Reductionist denies is that the subject of experiences is a *separately existing entity*, distinct from a brain and body, and a series of physical and mental events.<sup>12</sup>

It would not be true to say that there is *no* sense in which the Reductionist is an error theorist. To the extent that most of us believe—or so the Reductionist claims—that the referent of the word "I" is a Cartesian ego, we stand convicted of metaphysical error. But to deny that subjects of experience are *separately* existing entities is not to deny that subjects of experience or persons exist.<sup>13</sup> What the Reductionist aims to do is rather to give an account of what the existence of a person *consists in*; it consists, precisely, in the existence of his brain and body, the doing of his deeds, and the occurrence of his mental states and events.

If the no-subject theorist suggests that the idea of subjects of experience is a linguistically generated illusion, and that this ought to be conceded by the Reductionist, it is wholly inappropriate for

<sup>11</sup> P. F. Strawson, *Individuals* (London: Methuen, 1959), 94.

<sup>12</sup> Parfit, *Reasons and Persons*, 223.

<sup>13</sup> Parfit's clearest statement of this point is to be found in appendix D of *Reasons and Persons*.

the Reductionist to respond that there are subjects of experience because “this is the way in which we talk.” The fact that we ascribe experiences to ourselves in the form of subject-predicate propositions is, of course, acknowledged by the no-subject theorist. His point is precisely that surface grammatical considerations should *not* be taken as settling any metaphysical issues. In contrast, Parfit’s sole argument for embracing (5) rather than (4), and for claiming that people exist as subjects of experience, is a direct appeal to surface grammar. Not surprisingly, this leaves him open to the charge that his version of Reductionism only allows for the existence of subjects of experience as a figure of speech; we think and talk *as if* there are subjects of experience, distinct from series of mental and physical events, but, strictly speaking, these ways of thinking and talking might be said to embody a metaphysical error. The justification for our thinking and talking in this way cannot be that we *do* so think and talk.

Does this amount to the view that all forms of reductionism constitute error theories? Reductionism is sometimes understood as a theory of the *meaning* of propositions of a certain sort, and reductionism in this sense does often amount to a rejection of the area of discourse of which the reduction is proposed. But there is another form of reductionism—what might be called ontological reductionism—which is concerned with the nature of reality without being committed to any particular analysis of propositional content. For example, one kind of phenomenalist might argue that facts about the physical world are grounded in or sustained by facts about human sense-experience, without holding that propositions about physical objects are equivalent in meaning to propositions about human sense-experience.<sup>14</sup> Would such an ontologically reductive phenomenism amount to a repudiation of the physical world, or merely an explanation of what the existence of the physical world consists in? More generally, is there room for an ontological reductionism which is explicative without being eliminative?<sup>15</sup>

---

<sup>14</sup> John Foster emphasizes this distinction in his discussion of phenomenism in *The Case for Idealism* (London: Routledge and Kegan-Paul, 1982), 15–16.

<sup>15</sup> In asking this question, one is at odds with Quine’s view that the distinction between eliminative and explicative reductionism is unreal. See Quine’s discussion in *Word and Object* (Cambridge: The MIT Press, 1960), 265.

Suppose that the existence of entities of some given type  $T_1$ , or the occurrence of some given type of event, is held to consist in the existence of some apparently different type of entity  $T_2$ , or the occurrence of some apparently different type of event. Such a claim will be eliminative with respect to entities or events of the first type if it holds that references to them have no autonomous explanatory power, and that it suffices for an understanding of what an entity or event of this type *is* that one should grasp the fact that their existence just consists in the existence or occurrence of entities or events of type  $T_2$ . In contrast, a reductionist who *denies* these claims of explanatory redundancy will be noneliminative. For example, even if one holds that coming to believe is a neural event, one might maintain that one does not know what a belief *is* if one fails to recognize that beliefs are governed by what Davidson calls the “constitutive ideal of rationality,”<sup>16</sup> and that the explanation of action by the ascription of beliefs is illuminating to an extent which cannot be matched by explanation in purely neural terms. In the case of persons, a theory which accepts (3) may escape the charge of eliminativism by holding that the ascription of experiences to persons or subjects is essential for an adequate explanation of the unity of consciousness, and hence for a proper understanding of what a subject of experience *is*. If the unity of consciousness can only be explained by thinking of persons as subjects of experiences rather than a bundle of mental and physical events, then, to this extent, there will be a case for accepting (5) rather than (4).

What unites the different experiences had by a single person, either at a given time or at various times? Kant’s answer is that “all *my* representations in any given intuition must be subject to that condition under which alone I can ascribe them to the identical self as *my* representations, and so can comprehend them as synthetically combined in one apperception through the general expression ‘I think’ ” (B138).<sup>17</sup> Kant’s account of what unites diverse simultaneously or successively occurring experiences in a single

---

<sup>16</sup> Donald Davidson, “Mental Events,” in *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 223. My account of how an ontological reductionist might be noneliminative owes something to Davidson’s discussion, although Davidson rejects the characterization of his “anomalous monism” as a form of reductionism.

<sup>17</sup> All references in this form will be to the *Critique of Pure Reason*, trans. Norman Kemp Smith (London: MacMillan, 1929).

consciousness incorporates two elements: firstly, the subject must be capable of ascribing these experiences to himself as *his* experiences; and secondly, the subject must be aware, or potentially aware, of the numerical identity of that to which the different experiences are ascribed by him.

Does this not beg the question? Kant's view already assumes that experiences have subjects, and is concerned with the question of what unites the experiences of a *single* subject. Is it not what Kant assumes that is really at issue? The impression that Kant begs the question is misleading. Suppose that there is a series of experiences which are to be conceptualizable as experiences of independent, spatial objects. Kant's argument is that if one abstracts from the fact that experiences are states of a subject who can self-ascribe them and *think* of himself as their single subject, then the possibility of conceptualizing them in the manner described will not have been provided for. If diverse experiences are to be conceptualizable as glimpses of enduring objects, then they must belong to a subject who is capable of self-ascribing them. Given that Kant *defines* the unity of consciousness in terms of the possible self-ascription of experiences by an identical subject, another way of making the same point would be to say that the objective bearing of experiences requires the unity of consciousness in precisely the sense defined by him.

Parfit regards his Reductionism as incompatible with the Kantian position. He claims that if this position were the correct one, then it might refute his claim that our lives could be redescribed in an impersonal way. If the explanation of the unity of consciousness involves ascribing experiences to subjects, then a description of reality which fails to mention persons could not be complete.<sup>18</sup> Par-

---

<sup>18</sup> To avoid confusion, I will refer to any position which accepts Parfit's thesis (3) as "reductionist." Any position which accepts both (3) *and* Parfit's thesis (9), along with the corresponding impersonal account of the unity of consciousness, will be referred to as "Reductionist." Hence, Kant's position appears to be reductionist, without being Reductionist. I will also assume, from now on, that for a theory to qualify as a bundle theory, it must accept not only (3), but also thesis (4) rather than (5). This might prompt one to withdraw the introductory characterizations of Parfit's Reductionism as a version—even a highly sophisticated version—of a bundle theory of the self. However, if the argument of this paper is correct, then Parfit's right to (5) is very much open to question. For a suggestion that Reductionism and the Kantian account of the unity of consciousness are incompatible, see p. 225 of Parfit's discussion.

fit's response is, of course, to reject the Kantian account of the unity of consciousness. The unity of consciousness at a time is explained by the fact that several experiences can be co-conscious.<sup>19</sup> The transtemporal unity of consciousness is explained in terms of the R-relatedness of mental states, where R is the relation of psychological continuity or connectedness with any cause. Neither the R-relation nor the account of co-consciousness requires the ascription of experiences to a subject. We only need to "describe what, at different times, was thought and felt and observed and done, and how these various events were interrelated. Persons would be mentioned here only in the description of the *content* of many thoughts, desires, memories, and so on. Persons need not be claimed to be the thinkers of any of these thoughts."<sup>20</sup>

Suppose that Parfit is right about this, and that reference to subjects of experience is explanatorily redundant. How, in that case, could he meet the charge that his ontological Reductionism about the self is covertly eliminative? If appeals to the way in which we talk are beside the point, and the idea of subjects of experience does no work in the Reductionist's *explanatory* story, it is not clear how he can motivate the claim that there are subjects of experience, and that such subjects are distinct from the states ascribed to them. But perhaps Parfit should not have conceded that Reductionism and the Kantian position are incompatible. Suppose that the Kantian argument is sound, and that it provides a way of combining (3) and (5). Given that this combination is precisely what the Reductionist wants, why not regard Kant's argument as deepening the basic Reductionist insight rather than refuting it? A weakness of Reductionism is that it treats first-person thinking, the thinking of self-ascriptive thoughts, as a brute fact, without ever making clear the *point* of such thinking. Why should the Kantian argument not be regarded as making good this deficiency, without bringing with it any commitment to the existence of subjects as *separately* existing entities? If this reconciliation is to be successful, the appearance of conflict over the possibility of redescribing our lives in an impersonal way will need to be explained away.

The best way to proceed at this point would be to suggest that Kant and the Reductionist are operating at different levels, and, to

<sup>19</sup> Parfit discusses the unity of consciousness in sect. 88.

<sup>20</sup> Parfit, *Reasons and Persons*, 251.

a considerable extent, answering different questions. Suppose that we, as theorists, are concerned to describe a given mental life, and to explain what its unity consists in. We may acknowledge that the mental life under consideration includes the thinking of a great many self-ascriptive thoughts, and that such thoughts are an essential part of the framework required to sustain the idea that many of the experiences which form a part of that life constitute glimpses of independent spatial objects. None of this requires that we, as theorists viewing the life from the outside, need ascribe these self-ascriptive thoughts to subjects. We may continue to explain the unity of the life impersonally from our external standpoint, while acknowledging that the impersonal conception is not one which could be adopted from a standpoint within the life.

It might be argued that the Kantian argument slides over this difference in levels. When that argument claims that we cannot abstract from the fact that experiences are ascribable to an identical subject, the natural question is, from which perspective can we not abstract from this fact? Within a given mental life, diverse experiences are thought of as ascribable to one and the same subject; it must, indeed, be possible for the experiences within a given mental life to be accompanied by an “I think.” The great appeal of the Kantian account is that, if successful, it would explain the *point* of the requirement that the mental states within a given life be ascribable to a single subject. But what is required to make a mental life intelligible “from the inside” need not be what is required to characterize and explain a mental life from—to borrow a phrase of John McDowell’s—a “sideways on” perspective.<sup>21</sup> The question “what explains the unity of consciousness?” is a question asked from the sideways on perspective, and Kant’s mistake is to assume that the answer must be faithful to the conception which, from within a mental life, underlies the capacity to conceptualize experiences as experiences of the objective.

Whether or not this way of accommodating Kant’s insights within a Parfitian framework is successful is a question which will

---

<sup>21</sup> John McDowell introduces this notion in a discussion of rule-following, although he is highly critical of the idea that it is possible for us to view our rule-following practices from a sideways perspective. See his paper “Non-Cognitivism and Rule-Following,” in *Wittgenstein: to Follow a Rule*, ed. Holtzman and Leich (London: Routledge and Kegan Paul, 1981), esp. 150.

be pursued in the concluding section. Before then, however, it is important to gain some understanding of the *basis* of Kant's claim that—at least from a standpoint within a mental life—the capacity to conceptualize perceptions as perceptions of independent spatial objects rests upon the possible self-ascription of those perceptions by an identical subject.

## II

Kant's references to the possible consciousness of the numerical identity of the subject of diverse experiences need to be handled with caution, for Kant also argued that “no fixed and abiding self can present itself in [the] flux of inner appearances” (A107). Although the possible consciousness of the identity of the self lies at the heart of the deduction, in the “Paralogisms,” Kant goes on to point out that “the identity of the consciousness of myself at different times is only a formal condition of my thoughts and their coherence, and in no way proves the numerical identity of my subject” (A363). What is required is an account of the work done by the notion of the identity of the self, but one which respects the deflationary argument of the “Paralogisms.”

At times, Kant himself is inclined to explain matters in a way which suggests that the point about numerical identity is a relatively trivial one. At one point he writes that “I am conscious of the self as identical in respect of the manifold of representations that are given to me in an intuition, because I call them one and all *my* representations, and so apprehend them as constituting *one* intuition” (B135). Thus, if  $R_1$ ,  $R_2$  and  $R_3$  are all representations given to *me*, then it is indeed the numerically identical self which is conscious of these representations. This, however, is simply a consequence of the way in which the situation has been described; if I am conscious of various representations, then it is indeed the same self which is conscious of these representations. What becomes of the numerical identity of the self when the situation is described in less question-begging terms? If representation  $R_1$  occurs at  $t_1$  and is self-ascribed by its subject,  $R_2$  occurs at  $t_2$  and is self-ascribed by its subject, and  $R_3$  occurs at  $t_3$  and is self-ascribed by its subject, in each case the subject will think of itself as “I,” and will accompany its representations with an “I think,” but it does not follow that it is the same subject in each case. As Kant remarks, “Despite the

logical identity of the 'I', such a change may have occurred in it as does not allow of the retention of its identity, and yet we may ascribe to it the same-sounding 'I . . .' (A363). Knowledge of the identity of the subject must be based upon intuition or empirical awareness, and cannot be acquired through mere analysis of the proposition "I think." The dilemma, then, is that either the identity of the subject is built into the description of the situation, in which case the transition from the self-ascription of experiences to the identity of the subject of experiences will be uninteresting, or it is not built in, in which case it will simply be a mistake to conclude from the fact that all experiences must be self-ascribable that it is the same subject which is doing the experiencing and self-ascribing.

One way of bringing out the significance of Kant's talk of the numerical identity of the self would be to consider how the notion comes to be introduced during the discussion of the various forms of synthesis in the first-edition deduction. Kant's starting point is the claim that "if each representation were completely foreign to every other, standing apart in isolation, no such thing as knowledge would ever arise" (A97). Presumably, the point here is that for knowledge to be possible, representations must occur in regular sequences, so that the mind can discern patterns in experience and form expectations about future experience. In other words, not only must representations *be* orderly, the subject must also be aware of them *as* orderly and interconnected. This might explain Kant's definition of knowledge as essentially "a whole in which representations stand compared and connected" (A97). Understood in this way, knowledge involves both receptivity and spontaneity, both a passive and an active element. The active element is the capacity to synthesize representations. Kant speaks of all knowledge as involving a threefold synthesis, and at this point it becomes important to determine whether synthesis has to do with the *connectedness* of representations, or with the *awareness* of this connectedness, with that element of knowledge which involves the "comparison" of representations. As the discussion proceeds, and Kant's idealism becomes more prominent, it becomes apparent that he believes that the connectedness of representations is itself the product of a special kind of synthesis, namely, transcendental synthesis. Still, there is also a less extravagant conception of synthesis at play in the deduction, and this will be the focus of the present discussion.

The first stage of the threefold synthesis is the synthesis of apprehension or empirical synthesis. What is given to us in experience is “appearance,” and appearances are called perceptions when they are combined with consciousness. This act or process of becoming conscious of appearances is just what is referred to as the synthesis of apprehension (see A119-120 and B160). The second element of the threefold synthesis is the synthesis of reproduction in imagination. For Kant, “imagination” is a quasi-technical term; it is “the faculty of representing in intuition an object that is *not itself present*” (B151). Imagination is involved in the case in which representations of one type tend to occur in conjunction with representations of some other type, so that the occurrence of a representation of the first type results in a transition of the imagination to a representation of the second type. Imagination, in Kant’s sense, is also involved in what might more naturally be described as the operations of memory. To remember a previous perception is, in Kant’s terminology, to reproduce it in imagination. Indeed, Jonathan Bennett suggests that the Kantian notion of imagination might be identical with that of intellectually disciplined memory,<sup>22</sup> and while this proposal may not capture all the complexities of the Kantian notion, there can be little doubt that Kant regarded that notion of imagination which *is* closely connected with the concept of memory as being of enormous importance. It is in grasping the role of memory in perceptual representation that the importance of the idea of a numerically identical subject should eventually become apparent.

Kant’s thesis is that memory—the capacity to reproduce appearances—is a necessary condition of the possibility of experience, insofar as it is required for the formation of “complete” mental representations. As he puts it:

For experience as such necessarily presupposes the reproducibility of appearances. When I seek to draw a line in thought, or to think of the time from one noon to another, or even to represent to myself some particular number, obviously the various manifold representations that are involved must be apprehended by me in thought one after the other. But if I were always to drop out of thought the

---

<sup>22</sup> The suggestion is made in *Kant’s Analytic* (Cambridge: Cambridge University Press, 1966), 136.

preceding representations . . . and did not reproduce them while advancing to those that follow, a complete representation would never be obtained. (A101-102)

There is a shift here from talk about the role of memory in perception to talk about its role in thought, but the context suggests that the argument is intended to apply to both. In connection with perception, Kant's point may be illustrated by borrowing an example from the "Second Analogy," the example of someone perceiving a house. The idea is that an individual successively perceives different parts of a single house, with the successive perceptions building up or contributing to the picture of a single building whose various successively perceived parts are nevertheless conceived of as existing simultaneously. If, having perceived an upstairs window, one perceives the front door, one's perception of the front door will only contribute to the building up of a picture of a house of which the door is but one part if the preceding perception (of the upstairs window) is not forgotten. No doubt, Kant's way of stating the point is rather crude and liable to promote misunderstanding. His reference to the reproduction of preceding representations being necessary for the production of a complete representation suggests that one's perception of each part of the house contributes to the filling in of a mental jigsaw puzzle, with each successive perception being accompanied by a recollection of the preceding perceptions, until the mental picture is complete. Still, what does seem plausible is that in order to see what one sees *as* a part of a house, one must be able to retain information derived from preceding perceptions of other parts of the same building. Even if it is not literally a matter of picturing the upstairs window as one perceives the front door, one's perception of the door would not have the significance which it has if the preceding perception had simply "dropped out of thought."

Kant finds the example of the house attractive because of his insistence that the apprehension of the manifold of appearance is always successive (B234). But what of the case in which one perceives an object as a whole at a glance? It is not a matter of successively perceiving different parts of an object, so memory cannot be involved in this case as it was involved in the earlier case. Nevertheless, memory or the synthesis of reproduction *is* involved. Consider the case in which one perceives what one takes to be a single enduring object. Perceptions are transient, and, as Strawson has argued, we

could not count a single transient perception as a perception of an enduring, independent object “unless we were prepared or ready to count some different perceptions as perceptions of one and the same enduring and distinct object.”<sup>23</sup> In some cases these other perceptions will be merely possible perceptions, but in other cases—those in which the object is familiar—the other perceptions will include actual past perceptions of the same object. It is not a matter of recalling any particular past perception of a currently perceived object, but rather, of the present perception being “soaked with or animated by, or infused with . . . the thought of the other past or possible perceptions of the same object.”<sup>24</sup> For past perceptions to be intimately connected in this way with the present perception, those past perceptions must have been retained in memory. In Kant’s terminology, the subject must be capable of reproducing them.

The third element of the threefold synthesis is the synthesis of recognition in a concept. This element is introduced as underpinning the reproducibility of representations: “If we were not conscious that what we think is the same as what we thought a moment before, all reproduction in the series of representations would be useless” (A103). Kant’s point here is that in taking a transient perception to be a perception of an enduring object, say a tree, it is not enough that some other actual or possible tree-perception should be alive in the present perception. It is also required that the other perception be *conceptualized* or *recognized* as a perception of the *very same* tree as the tree of which current perception is a perception. The idea being captured here is that there are or could be different perceptions of the very same thing. In fact, this is a point which Kant has already made in connection with the synthesis of reproduction, but he deems the capacity to conceptualize diverse perceptions as perceptions of the same object to be of sufficient importance to merit a separate title, the “synthesis of recognition in a concept.”

Kant’s theory of synthesis was introduced into the discussion in the hope that it would explain why the unity of consciousness was held to require the possible self-ascription of experiences by a subject who is capable of grasping the idea that it is one and the

---

<sup>23</sup> P. F. Strawson, “Imagination and Perception,” in *Freedom and Resentment and Other Essays* (London: Methuen, 1974), 52.

<sup>24</sup> *ibid.*, 53.

same subject to whom the different experiences are being ascribed. So far, however, nothing has been said in defense or explanation of this conception of unity of consciousness. Still, it would be a mistake to conclude that the discussion of the threefold synthesis has been an unnecessary detour, for as Kant's discussion proceeds, material from the theory of synthesis turns out to have an important bearing upon the defense of the theory of the unity of consciousness. The key to understanding the relationship between the different elements of Kant's account is the notion of experience of independent objects. It will have been noted that Kant's theory of synthesis, as here interpreted, simply assumes that perceptions are conceptualized as perceptions of independent, enduring objects. In fact, Kant's view was that it is possible to *argue* for an intimate connection between the idea of experience of independent objects and the unity of consciousness. On the one hand, experience of objects is held to be required for the unity of consciousness. On the other hand, it is also argued that for representations to count as representations of objects, they must belong to a unified consciousness. As Kant puts it, "There can be in us no modes of knowledge, no connection or unity of one mode of knowledge with another, without that unity of consciousness which precedes all data of intuitions, and by relation to which representation of objects is alone possible" (A107). The question which is now to be considered is *why* the unity of consciousness is held to be necessary for the representation of objects. It is in understanding Kant's answer to this question that the connection between the various elements of the discussion so far will become apparent, for Kant's answer ties together a great many of the most important strands of the deduction.

Kant's thesis is clearly stated in the following important passage from the first-edition deduction:

Intuitions are nothing to us, and do not in the least concern us if they cannot be taken up into consciousness, in which they may participate either directly or indirectly. In this way alone is any knowledge possible. We are conscious *a priori* of the complete identity of the self in respect of all representations which can ever belong to our knowledge, as being a necessary condition of the possibility of all representations. For in me they can represent something only insofar as they belong with all others to one consciousness, and therefore must be at least capable of being so connected. (A116)

Taken in the context of Kant's concern with the notion of experience of the objective in the deduction, what this suggests is that for in-

tutions to be conceptualizable as representations of an objective world, they must be self-ascribable by a subject capable of consciousness of his own identity as the single subject of the representations. Unfortunately, it is far from clear how this thesis is to be defended. If a subject self-ascribes a perception P and is prepared to count other perceptions as perceptions of the same thing as the thing of which P is a perception, surely there is no requirement that he also think of the other perceptions as *his*. Imagination reproduces absent perceptions, and such perceptions may need to be recognizable as perceptions of the same object, but why is there a connection between the identity of the *object* of the perceptions and the identity of their subject? If it is said that imagination can only reproduce *my* past perceptions, and that this is the crux of the argument, it will no longer be clear what *special* work is done by the idea of experience of an objective world.

In *The Bounds of Sense*, Strawson puts forward an argument which—perhaps unintentionally—sharply illustrates the difficulties with which Kant’s argument has to contend. Strawson raises the question of what it is for a series of experiences to build up the picture of an objective world. His answer is that a series of experiences capable of playing this part has a double aspect:

On the one hand it cumulatively builds up a picture of a world in which objects and happenings (with their particular characteristics) are presented as possessing an objective order, an order which is logically independent of any particular experiential route through the world. On the other hand it possesses its own order as a series of experiences of objects. If one thought of such a series of experiences as continuously articulated in a series of judgements, then, taking their order and content together, those judgements would be such as to yield, on the one hand, a (potential) description of an objective world and on the other the chart of the course of a single subjective experience of that world.<sup>25</sup>

To talk of a single, subjective, experiential route through the world is not to talk of the route through the world traced by a person. It is possible to speak of experiences as constituting a subjective route through or temporally extended point of view on the world while abstracting from the fact that these are experiences ascribable to a subject of experience. The fact that experiences are conceptual-

---

<sup>25</sup> P. F. Strawson, *The Bounds of Sense* (London: Methuen, 1966), 105–106.

izable as constituting a single point of view on the world is represented as a necessary but not a sufficient condition for the ascription of experiences to a subject. This feature of Strawson's position enables one to do justice to two important features of Kant's position. On the one hand, the thesis that representations can only function as representations of the objective if they are combined in a single consciousness which is aware of its numerical identity through various representations may now be understood as the thesis that for a series of representations to count as representations of the objective, they must be conceptualizable as constituting *one* subjective route through the world. On the other hand, given that the notion of a single subjective route through or temporally extended point of view on the world is an abstraction from the empirically applicable concept of a subject of experience, the fact that experiences are conceptualizable as constituting an experiential route is not sufficient to show that those experiences are experiences of a single subject or person. The need for representations of the objective to belong to one consciousness—that is, to constitute a single point of view—is in no way sufficient to prove that there is a single, persisting subject of experience.

Strawson's argument may enable one to reconcile the "objectivity requires unity of consciousness" argument of the "Transcendental Deduction" with the deflationary argument of the "Paralogisms," but it also plays into the hands of the bundle theorist. Far from troubling the bundle theorist, the notion of an experiential route through the world will simply provide him with further ammunition, in so far as this notion is an abstraction from the full-fledged notion of a subject of experience. Indeed, one way of understanding the bundle theorists' position would be to understand it as, precisely, the view that the notion of a single, subjective, experiential route through an objective world *can* stand on its own, without ultimately needing to rely upon the possible ascription of the experiences which combine to constitute such a route to a persisting *subject* of experience. No doubt, the Kantian will not see the notion of an experiential route as ultimately viable independently of the notion of a subject to whom a series of experiences can be ascribed, but what is now called for is an argument against the possibility of independent viability.

One way to work up to a defense of Kant's "objectivity requires unity of consciousness" argument would be to begin with the observation that being able to conceptualize perceptions as perceptions

of enduring objects requires the capacity to apply the distinction between having successive perceptions of one and the same object, and successively perceiving qualitatively similar but numerically distinct objects. Thus, consider a case in which there are distinct perceptions ( $P_1$ ,  $P_2$ ) which are to be conceptualized as successive perceptions of a single enduring object  $O_1$ . However, the environment may contain more than one O-type object. Given that there may be another object  $O_2$  which is qualitatively indistinguishable from  $O_1$ , on what basis are  $P_1$  and  $P_2$  to be conceptualized as successive perceptions of  $O_1$  rather than of  $O_2$ , or perceptions first of  $O_1$  and then of  $O_2$ ? One answer to this difficulty is suggested by the simple theory of perception which is implicit in ordinary thought about the world and our place in it. According to the simple theory, perceiving some object or phenomenon requires that one be appropriately located—both spatially and temporally—with respect to that object or phenomenon. This idea of the enabling conditions of perception provides for the idea of existence unperceived, for the fact that a perceivable object may not actually be perceived may now be accounted for in terms of the enabling conditions of perception not being met.<sup>26</sup> The simple theory of perception also provides an explanation of the basis upon which distinct perceptions may be pinned down as perceptions of one particular object  $O_1$  rather than some spatially distant duplicate, because given the location and orientation of the perceiver,  $O_1$  was the only O-type object perceivable at the times at which the perceptions  $O_1$  and  $O_2$  occurred. If the perceiver has the conception of himself as appropriately located with respect to  $O_1$ , and not to any other O-type object, then he has the resources to conceptualize his successive perceptions as successive perceptions of the single object  $O_1$ .

This argument has a bearing upon the earlier discussion of Kant's theory of synthesis. According to that theory, for a given transient perception to be conceptualizable as a glimpse of an enduring object, one must be prepared to count other perceptions as perceptions of the very same object. The bringing to bear of other

---

<sup>26</sup> The connection between the simple theory of perception and the idea of existence unperceived is discussed by Gareth Evans in part two of his "Things Without the Mind—A Commentary upon Chapter Two of Strawson's *Individuals*," in *Philosophical Subjects: Essays presented to P. F. Strawson*, ed. Zak van Straaten (Oxford: Clarendon Press, 1980), 88–90.

perceptions upon a current perception is the synthesis of reproduction in imagination; the interpretation of the reproduced perceptions as perceptions of the same object as the currently perceived object is the synthesis of recognition in a concept. There will be cases in which the perceived object is one which has been previously perceived, perhaps just a few moments ago, and in which the other perceptions of the currently perceived object which are alive in the present perception are actual past perceptions of that object, and not merely imaginary perceptions. What is involved in the thought that the recalled or “reproduced” perceptions were perceptions of the very same object? So far in the account of the synthesis of recognition in a concept, the subject’s capacity to think of past perceptions as perceptions of the very same object as that which is currently being perceived has been taken for granted. It should be acknowledged, however, that such thoughts cannot stand in isolation and that they need to be anchored in some way. It is not enough to speak of the subject as having the capacity to think of or conceptualize his perceptions in a certain way without explaining what is involved in the subject’s possession of this capacity. For example, a subject with this conceptual capacity must be able to distinguish between the case in which a reproduced perception may be counted as a perception of *this* very object (the one currently being perceived) and the case in which the reproduced perception is a perception of a qualitatively identical but numerically distinct object. It is in this connection that the ability to ascribe the reproduced perception to the very same *subject* as the subject of the current perceptual experience assumes considerable significance.

Earlier it was suggested that there is a simple answer to the question why a recalled perception needs to be thought of as one’s *own* earlier perception, for it might be held to be a conceptual truth that one can only remember or “reproduce” one’s own past experiences. It was objected that this did not explain the *special* connection between thinking of past perceptions as one’s own and thinking of perceptions as glimpses of enduring objects, but it is now possible to see the connection. Suppose that a current perception,  $P_2$ , and a remembered perception,  $P_1$ , are to be conceptualized as perceptions of  $O_1$ . A simple theory of perception was appealed to earlier in explaining the basis upon which a current perception may be determined to be a perception of a given object rather than a spatially

distant duplicate, and the same simple theory may be combined with the ordinary concept of memory to explain the basis upon which a remembered perception may be conceptualized as a past perception of  $O_1$  rather than  $O_2$ . My own past perceptual experiences are the only ones which I can correctly be said to remember, and only objects with respect to which I am appropriately located can be perceived by me. So a remembered perception  $P_1$  can be conceptualized as a perception of  $O_1$  rather than  $O_2$  just if my route through the world was such that at the time of  $P_1$ 's occurrence, I was appropriately located with respect to  $O_1$  rather than  $O_2$ . For example, if a replica of Marble Arch were built in Dallas, I could assure myself that the Marble Arch which I remember seeing was the one in London by appealing to the fact that I have been to London but not to Dallas. If I did not think of the earlier perception as *mine* ("ascribe it to the identical subject"), the fact that *I* have been to London but not to Dallas would no longer help to sustain the thought that the perception was a perception of the building in London rather than the one in Dallas. My own location at the time of  $P_1$ 's occurrence simply drops out of consideration as irrelevant unless I ascribe  $P_1$  to myself.

This argument relies upon a conception of the enabling conditions of memory which has been well described by Shoemaker:

The spatiotemporal region which is "rememberable" by a given person can be charted by specifying the intervals of past time during which the person was conscious, and by specifying the person's spatial location, and indicating what portions of his environment he was in a position to witness, at each moment during those intervals. If someone reports that he remembers an event of a certain kind, we know that unless his memory is mistaken an event of that kind occurred within the spatiotemporal region rememberable by him, and in principle we can chart this region by tracing his history back to its beginning.<sup>27</sup>

In effect, the earlier argument appealed to this account of the enabling conditions of memory in explaining the Kantian idea that there is an intimate connection between our way of making sense of the numerical identity of the objects of diverse perceptual experiences and the identity of the subjects of those experiences.

---

<sup>27</sup> Sydney Shoemaker, "Persons and their Pasts," in *Identity, Cause, and Mind* (Cambridge: Cambridge University Press, 1984), 28. I am much indebted to this important paper.

Successive perceptions are conceptualizable as perceptions of the very same thing precisely because of the possibility of ascribing those perceptions to a numerically identical subject whose route through the world anchors those perceptions to one particular object. It is not enough to say that it must be possible for a current perception to be soaked by the thought of a past perception of the very same object. The identity claim needs to be anchored, and if the earlier argument was correct, then it is the numerical identity of the subject of diverse experiences which provides the anchoring. And for the numerical identity of the subject to play this part, those experiences must be thought of as having a subject.

This argument might also be held to provide an answer to the question why the Strawsonian notion of a single, subjective experiential route through an objective world should not ultimately be viable independent of the idea that the experiences which constitute such a route are ascribable to an identical subject of those experiences. The notion of an experiential route was introduced as an abstraction from the rich conception of a subject of experience, but the preceding argument suggests that the two cannot be independent of one another. To introduce the notion of an experiential route is to invite questions about the criteria of identity for such routes; it is to invite questions about how distinct experiences are to be determined to belong to and be part of one experiential route rather than two distinct routes. If all one is entitled to say is that first one O-type perception occurred, and then another, it is, as yet, far from clear what the force is of either asserting or denying that those experiences are elements of a single route, and without the idea of a single route to rely upon, the idea that perceptions are perceptions of a single object is rendered similarly problematic. Once the idea that given experiences are episodes in the history of a single *person* is in place, and once experiential routes are seen to be based upon personal routes, these difficulties become far less intractable. Diverse experiences may be conceptualized as elements of a single, subjective, experiential route because of the possibility of appealing to the notion of an enduring subject of those experiences, a physical thing located in space. As Strawson himself later remarks, "Instead of talking, dubiously, of an experiential route through the world, we may talk, confidently, of an undeniably persistent object, a man, who perceptibly traces a physical, spatiotemporal route through the world and to whom a series of experiences may be ascribed with no

fear that there is nothing persistent to which they are being ascribed.”<sup>28</sup>

Is the preceding attempt to construct an “objectivity requires unity of consciousness” argument on Kant’s behalf compatible with the deflationary remarks of the “Paralogism?” In a famous footnote (A364), Kant describes an example in which what the Cartesian regards as the life of a single soul in fact involves the existence of a series of souls, with the states of consciousness of each member of the series being passed along to the next member. A more modern version of such a fantasy would be a case in which, as a result of a partial brain transplant, one person’s memories are transferred into some other person’s brain. In such cases, its seeming to the recipient that he saw F at some earlier time may be the result of someone else’s having seen F at that time. So if the recipient were to conceptualize the earlier experience as *his* experience, he would simply be mistaken. Even if a subject *thinks* of experiences in a certain way—that is, ascribes them to himself—it does not follow that he has *knowledge* of his own identity as the subject of these experiences. As Kant himself writes, “I do not know an object merely in that I think, but only in so far as I determine a given intuition with respect to the unity of consciousness in which all thought consists” (B406).

Before proceeding any further, it would be as well to register one cautionary remark. It has been claimed that the earlier defense of the “objectivity requires unity” thesis is at least Kantian in spirit, but it must be conceded that it is not the defense which Kant himself proposed in the “Deduction.” Kant certainly asserts a connection between the idea of experience of independent objects and the unity of consciousness, but he barely argues for this thesis. What has been proposed here is simply one way of arguing for the thesis, but the temptation to claim that it might be what Kant “had in mind” ought to be resisted. The proposed defense turns on the notion of the subject as a physical thing tracing a path through space, but this is a notion of which Kant makes nothing in the “Deduction.” This is a surprising omission, but the important point is that no attempt should be made to exaggerate the faithfulness of the argument to Kant’s own record.

Leaving aside questions of exegesis, what now needs to be considered is the objection that, whatever its origin, the proposed de-

---

<sup>28</sup> Strawson, *The Bounds of Sense*, 164.

fense of Kant's thesis is in fact unsuccessful. The objection is that the defense is committed to an unacceptable conceptual foundationalism in which the location of the subject is always taken for granted in determining the identity of the objects of perception.<sup>29</sup> What this assumption ignores is the obvious fact that just as one might argue from where one is to what it is that one is perceiving, so one might also infer one's location from what one can see. To return to an earlier example, I might assure myself that what I am seeing is the original Marble Arch rather than an indistinguishable duplicate by appealing to the fact that I am in London, but my grounds for believing that I am in London will surely include the assumption that what I can see is, indeed, Marble Arch. When pressed, the argument rapidly leads to a vicious circularity. In the next section, this objection will be considered in some detail and the defense of the Kantian thesis refined in various ways. As will become apparent, the necessary refinements are not such as to call into question the validity of the basic Kantian insight.

### III

Suppose that as I look at London's Marble Arch, I seem to remember an earlier perception,  $P_1$ , as of a Marble Arch-type monument. Was  $P_1$  a perception of the very same monument as the monument of which the current perception,  $P_4$ , is a perception? The circularity objection at the end of the last section assumed that this question is to be answered upon the basis of  $P_1$  and  $P_4$  alone, but this is clearly not the case, for one might also appeal to the intervening perceptions  $P_2$  and  $P_3$ . Suppose that  $P_2$  was as of Hyde Park Corner and  $P_3$  as of Park Lane, and that the time elapsed from  $P_1$  to  $P_4$  is no more than a few minutes. Given my knowledge of the fact that in London, Marble Arch, Hyde Park Corner, and Park Lane are spatially adjacent places, I have grounds for anchoring  $P_1$  and  $P_4$  to the very same monument assuming that all the perceptions are ascribed to a single subject, namely myself. The point here is, as Evans puts it, that "self-location cannot in general be a momentary thing."<sup>30</sup> That is, the subject must have the capacity to "retain

<sup>29</sup> John Campbell put this objection to me.

<sup>30</sup> Gareth Evans, *The Varieties of Reference* (Oxford: Oxford University Press, 1982), 243.

information about his previous perceptions, and to use that information in making judgements about his past, and thereby his present position.”<sup>31</sup> If  $P_1$  were to be conceptualized or treated as a perception of Marble Arch’s Texan duplicate, the perceptions  $P_1$  and  $P_4$  would no longer be explainable by reference to the idea of their subject travelling through adjacent places in London. Given that the Texan duplicate’s surroundings do not include anything like Park Lane and Hyde Park Corner, treating  $P_1$  as a perception of the duplicate would require the assumption of one’s having travelled a very considerable distance within a short space of time without one’s knowing it. The point is not that this could not have happened, but that it is an *assumption* of our theory of self-location that such things do not generally happen.

Following Christopher Peacocke’s terminology in *Holistic Explanation*, the considerations of the last paragraph may be characterized as constituting *intertemporal restrictions* upon self-location. There are two related points here: the first is that one’s location is determined upon the basis of *sequences* of perceptions, where these sequences are in turn interpreted as relating to spatially *adjacent* places. In Peacocke’s words, “The spatial scheme constrains the possible sequences of experiences it explains to those corresponding to *paths* through the space.”<sup>32</sup> The second related restriction concerns the speed of one’s movement. Peacocke writes: “It could be argued that it is *a priori* that there is some restriction . . . on which places are accessible to the experiencer at a given time interval after he has been at a given place . . . if it were not so, the experiencer could not have empirical reasons for believing he is at one rather than another of two qualitatively similar places.”<sup>33</sup> Given these intertemporal restrictions, and some elementary geographical knowledge, questions about one’s past and present location and the relation between successively perceived objects can be answered simultaneously, without positing any objectionable conceptual priority. Given the intertemporal restrictions, the best explanation of the perceptions  $P_1$ – $P_4$  in the earlier example would be that they relate to adjacent places in London, and this in turn gives one

---

<sup>31</sup> Evans, *The Varieties of Reference*, 243.

<sup>32</sup> Christopher Peacocke, *Holistic Explanation: Action, Space, Interpretation* (Oxford: Clarendon Press, 1979), 33.

<sup>33</sup> *Ibid.*

grounds for treating  $P_1$  and  $P_4$  as perceptions of the very same London monument rather than as perceptions of a monument in Dallas followed by a perception of a monument in London, or, for that matter, as perceptions of the very same Texan monument.

Moreover, the fundamental Kantian insight remains unchallenged, for the application of the intertemporal restriction requires that the successive perceptions to be explained by appeal to the restrictions all be ascribed to one and the same subject. In the example, suppose that, as the subject of  $P_4$ , I refuse to ascribe  $P_1$ – $P_3$  to myself. In that case, they will have no bearing upon *my* location, and the question of the relationship between the object of  $P_1$  and the object of  $P_4$  will once again be rendered problematic. Without being unified by means of their ascription to a single subject, the perceptions would have to be viewed as independent units; no single perception would have any bearing upon the interpretation of preceding and succeeding perceptions. The fact that  $P_2$  was a perception as of Hyde Park Corner *only* contributes to the interpretation of  $P_1$  as a perception of London's Marble Arch if  $P_1$  and  $P_2$  are ascribed to the same subject. Without this assumption, the appeal to how far a *single* subject could or could not have travelled within a short space of time would be quite ineffective, for the subject  $S_1$  of  $P_1$  may indeed have been located in Dallas and the subject  $S_2$  of  $P_2$  located in London, as long as it is not assumed that  $S_1 = S_2$ .

The moral of this discussion, then, is that perceptions are anchored to particular objects not by a *direct* appeal to the location of their subject, for the latter cannot simply be assumed. Rather, perceptions are anchored to specific objects and locations simultaneously, *via* their ascription to a single subject, in conjunction with the intertemporal restrictions. The point of this argument is not to deny that qualitative considerations have a significant part to play in the reidentification of the objects of perception. One skyscraper may, in certain respects, be very much like another, but it is rarely, if ever, the case that such buildings are literally qualitatively indistinguishable. So in determining whether or not a currently perceived building is the same as a previously perceived building, a subject will not, in general, be *confined* to spatial considerations. Nevertheless, spatial considerations do and must make a significant contribution to reidentification. Exclusive reliance upon qualitative considerations would require possession of a capacity to retain vast amounts of detailed information about the fea-

tures of perceived objects. Even if it is assumed that massive reduplication is not a predicament which we actually face, spatial considerations will not be deprived of their significance. In any case, it is not merely a matter of the greater economy and efficiency of such considerations; there is an underlying conceptual point here, namely that, as Evans puts it, "it is precisely what distinguishes individual-recognition from kind-recognition that the former is sensitive to considerations bearing upon the identity of a single object from time to time, and this means sensitive to spatio-temporal considerations."<sup>34</sup>

Still, there remain some difficulties with the refined Kantian argument, with its emphasis upon the intertemporal restrictions and the unity of consciousness, that is, the possible self-ascription of perceptions by one and the same subject. The first concerns the problem of circularity, for it might be argued that the recently introduced refinements do not really meet the difficulty. Consider the Twin Earth fantasy, in which it is supposed that somewhere in a distant galaxy is a planet which is qualitatively indistinguishable from Earth. In the earlier example, the fact that the perceptions  $P_1$  to  $P_4$  occurred in the order described was, in conjunction with the intertemporal restrictions, taken as grounds for supposing their subject to have been located in London. Given the Twin Earth hypothesis, however, it appears that there is nothing to rule out the possibility of their being successive perceptions of adjacent places in Twin Earth London. Given that in Twin London, duplicate Marble Arch does indeed lie in between duplicate Park Lane and duplicate Hyde Park Corner, the restrictions will not suffice to rule out the sceptical hypothesis that one is radically mistaken about one's own location.

In "Persons and their Pasts," Shoemaker responds as follows to the Twin Earth fantasy: if some person, S, reports having witnessed an explosion in a New York-like city, we are entitled to conclude that what he remembers witnessing took place in Earthly New York rather than some distant duplicate city since "what we can and do know is that New York is not reduplicated within any spatiotemporal region of which anyone with whom we converse can have had experience."<sup>35</sup> One difficulty with this response, and indeed

<sup>34</sup> Evans, *The Varieties of Reference*, 278.

<sup>35</sup> Shoemaker, "Persons and their Pasts," 29–30.

Shoemaker's discussion as a whole, is that it is concerned with the interpretation of testimony evidence, and this only serves to reproduce the original difficulty at another level. If the recipient of the information, R, already knows that he is on Earth, then he may well be entitled to interpret S's report as relating to events on Earth, given that S could not have travelled from Twin Earth to Earth within the time available, and given also that S's report is based upon memories of his own past experiences, and not quasi-memories of someone else's past experiences. But R's knowledge of his *own* location remains unexplained in Shoemaker's discussion. The point is that the intertemporal restrictions require a fixed point from which to work, either backwards or forwards; if I know where I am now, I may be able to work out where I was five minutes ago, and vice versa, but this does not seem to meet the *radical* sceptical point, which concerns the antecedent of this conditional.

A better response to radical, Twin Earth-based scepticism would be to suggest that it is, in a sense, irrefutable, but that it has no bearing upon the Kantian argument, which was concerned with the problem of reidentification. Even if central London is reduplicated on Twin Earth, it is unique "around here." If it is assumed that I could not have travelled millions of miles in the space of a few minutes without my knowing it, I may conclude that the Marble Arch-type monument which I can now see and the one which I saw five minutes ago could not have been distinct monuments, separated by millions of miles. The point is that "around here" is to be understood egocentrically, as indicating an area centering on the subject, and as long as there is only one Marble Arch-type monument around here, the possibility of massive reduplication does not undermine one's capacity to determine the currently perceived monument to be the very same (numerically) as a monument perceived previously. The claim that there is an egocentric element in reidentification need not be discussed at length here, for it has been lucidly argued for by Evans, in his discussion of recognition-based identification. Reidentification or recognition involves operating within an egocentrically identified "area of search," such that massive reduplication beyond the area of search does not undermine our possession of uniquely targeted recognitional capacities within the area. And, as Evans adds, "it is essential, for our possession of these absolutely discriminating capacities, that we move through space in regular and predictable ways, and that we are generally able to tell when

we move—as, indeed, we are, because movement, when we are conscious, has characteristics concomitant, and it does not generally take place (unless we have specifically arranged for it) during sleep.”<sup>36</sup>

Henceforth, it will be assumed that the circularity objection which was introduced at the end of section 2 can be defused. What now needs to be examined is an objection to the refined Kantian argument which is rather more closely related to the concerns of section 1, to the discussion of the plausibility of versions of the bundle theory.

The key to the refined Kantian argument was the claim that the application of the intertemporal restrictions requires that the successive perceptions to be explained by appeal to the restrictions all be ascribed to one and the same subject; as Kant might have put it, they must be taken up into a unified consciousness which can think its identity as the subject of the successive perceptions. This claim might be disputed. Suppose that two subjects, A and B, visit London. A regularly visits London, but it is B’s first visit. Neither has ever been to Dallas. As B sees Marble Arch, she has the powerful impression of having seen it before; she actually seems to remember a past perception of just such a building. Suppose also that prior to their visit, B had had some of A’s memory traces copied in her brain. What should B make of her apparent memory? As she has been to neither Dallas nor London before, one response would be for her to dismiss it as delusive, as a case of *déjà-vu*. Given her knowledge of the memory-trace transfer, however, would it not be far more plausible for B to suppose that she is in fact quasi-remembering one of A’s past experiences?<sup>37</sup> Since B knows that A has been to London before, but never to Dallas, would B not also be entitled to think of the quasi-remembered past perception as a perception of this very monument, the one she is currently perceiving, rather than the Texan duplicate? Here, then, is a case in which a subject can conceptualize a pair of perceptions as perceptions of one and the same thing without ascribing both perceptions to herself.

As it stands, this counterexample may not seem very threatening, for it is heavily dependent upon B’s access to extraneous in-

<sup>36</sup> Evans, *The Varieties of Reference*, 281.

<sup>37</sup> The example is a variation on Parfit’s Venetian Memories example in sect. 80 of *Reasons and Persons*.

formation of a sort which may not generally be available. Crucially, B must know of the memory trace transfer, and must also have independent knowledge of the fact that A has been to London before, but not to Dallas. The first of these points is a reflection of a more general feature of the epistemology of quasi-memory, which is that such memories can only yield knowledge about past experiences if one knows roughly how the quasi-memories have been caused.<sup>38</sup> The force of the counterexample is thus somewhat undermined, for it is difficult to believe that our capacity to recognize objects rests upon our possession of such extraneous information.

A better counterexample would run as follows: suppose that B seems to remember not just a past perception as of a Marble Arch-type monument, but rather a coherent sequence of perceptions, including perceptions as of Park Lane and Hyde Park Corner. Would B not be entitled to take it that these past perceptions were of London, and hence that the perception as of Marble Arch was of London's monument rather than its Texan double? Moreover, it might be argued, B need *not* think of them as her *own* past perceptions. B could reason that the subject of the past perceptions, *whoever* he or she was, must have been located in London, given the order and content of the perceptions, and the intertemporal restrictions. This is just the kind of reasoning that B engages in when she is dealing with what she takes to be her own past perceptions, so why should she not engage in parallel reasoning to fix the location of the subject of a series of past perceptions when she does not think of that subject as herself? The intertemporal restrictions are still applicable *within* the series of past perceptions, whether or not B ascribes those perceptions to herself. This might, of course, prompt the response that if B does not think of the past perceptions as her own, then, in the absence of extraneous information about the origin of her apparent memories of them, it would simply be a leap in the dark for B to suppose that she is in fact quasi-remembering some other subject's experiences rather than imagining things. This is true, but it might be wondered whether some reliance upon extraneous information is really all that problematic. Suppose that it is natural for us to quasi-remember some of our mothers' experiences, and to seem to remember them "from the inside." Upon B's first visit to

---

<sup>38</sup> Parfit, *Reasons and Persons*, 221.

London, she seems to remember a coherent set of past experiences as of London. It might be perfectly reasonable for her to suppose that she is quasi-remembering some of her mother's experiences, even if she does not have independent assurance that she had ever been to London. The apparently remembered perceptions *themselves* give B grounds for supposing that their subject, presumably her mother, had visited London, unless B has any evidence to the contrary. B does not *first* settle the question of the location of the subject of the past perceptions, and then anchor them to specific objects. Rather, the order and content of the perceptions themselves enable B to anchor them to specific objects, and their subject to specific locations, even if she does not think of herself as their subject.

What this example suggests is that Kant exaggerates the connection between the possibility of conceptualizing past and present perceptions as perceptions of one and the same object and their ascription to one and the same subject. Nevertheless, it should be noted that in the example, even if B does not ascribe the past perceptions to herself, she must, if the intertemporal constraints are to be applicable, ascribe them to a *single* subject. The ascription of past experiences to a single subject remains a part of the story. It might also be argued that even if in particular cases B might suppose that she is quasi-remembering someone else's experiences, she must *in general* think of apparently remembered past experiences as *her* own past experiences, if she is to be able to think of such experiences as contributing to a coherent picture of an objective world. This is not a matter which will be pursued. In the first place, the inconclusiveness of transcendental arguments, of attempts to establish strictly necessary conditions, ought to be conceded by the Kantian.<sup>39</sup> If facts of nature were very different from what they actually are, then perhaps our conceptual practices would be different, and we would have some other way of anchoring the distinction between the qualitative and numerical identity of the objects of distinct perceptions. What can be said is that the self-ascription of experiences by an identical subject constitutes *our* way of anchoring the distinction. In the second place, as noted earlier, the

---

<sup>39</sup> For a discussion of this point, see my "Transcendental Arguments, Transcendental Synthesis, and Transcendental Idealism," *The Philosophical Quarterly* 37 (1987), esp. the concluding section.

best Reductionist response to the Kantian argument would not be to challenge it on points of detail, but rather to attempt to accommodate it within the Reductionist picture, and to present it as deepening the insights of that picture rather than repudiating them. The question of whether or not this is a tenable response will be considered in the concluding section.

#### IV

The question which prompted the discussion of the last two sections was that of how an anti-Cartesian, reductionist view of the self might be entitled to the claim that there are subjects of experience, distinct from series of mental and physical events, while maintaining that the existence of a subject just consists in the existence of a brain and body, and the occurrence of a series of inter-related physical and mental events. The Kantian response was to suggest that the existence of subjects, to whom mental states are properly ascribable, played a crucial role in the explanation of the unity of consciousness, and in the patterns of reasoning required to sustain the idea of experience of spatial objects. To this extent, the Kantian reductionist argues, the existence of subjects of experience must be more than a linguistic illusion, a dispensable by-product of the way in which we talk. Parfit is reluctant to accept this because he thinks that it undermines the claim that it is possible to give a complete description of reality in impersonal terms. The suggestion was then made that it might be possible to reconcile the two positions by arguing that they operate at different levels. Now that the details of the Kantian argument have been presented, the proposed reconciliation may be assessed.

Some form of level distinction is very much in evidence in Parfit's discussion of the Cartesian Cogito argument. Unlike Georg Lichtenberg, Parfit allows that an adequate substitute for the Cogito need not be wholly impersonal. It needs to be recognized that there *are* "I"-thoughts, so instead of saying, "I think, therefore I am," Descartes should have said, "It is thought: I am thinking." The existence of a Cartesian subject cannot be deduced from this, for, "Since the subject of experiences here mentioned only is the *content* of the thought, this sentence does not ascribe the thought to a thinker."<sup>40</sup> In other words, the theorist attempting to describe and

<sup>40</sup> Parfit, *Reasons and Persons*, 221.

explain a mental life from the outside ought to acknowledge that there are self-ascriptive thoughts without *himself* ascribing such thoughts to subjects. By the same token, apparent memories may come to us in the first-person mode of presentation, but the concept of quasi-memory does not presuppose that of personal identity, for the theorist can continue to insist that for S's quasi-memory to be *veridical*, it need not be causally dependent upon S's own past experiences.

What these points suggest is that the Reductionist is committed to what might be called a "dual perspective" approach. The idea seems to be that the Reductionist can detach himself from a given mental life, and, while acknowledging the distinctive first-person content of many of the thoughts which form a part of the life, he may describe and explain the unity of the life in purely impersonal terms. As far as Kant's argument is concerned, the Reductionist response ought to be that Kant is guilty of failing to separate the answers to the following two quite separate questions: (a) What, from a standpoint within a given mental life, underpins the idea that that life includes experiences of independent spatial objects? (b) What does the unity of consciousness, of a given mental life, consist in? Kant's answer to the first question is, roughly, that the life must *also* include the thinking of self-ascriptive thoughts, and the belief that there is a single, persisting subject to whom mental states are properly ascribable. This is both true and important; the Reductionist ought to concede that the simple theory of self-location which is an essential element in spatial thinking does indeed require the *self*-ascription of past and present experiences by and to one and the same subject. So far, the Reductionist may add, there is nothing in the argument with which he is obliged to disagree. The Kantian only thinks that there is a difficulty here because he *defines* the unity of consciousness in terms of the possibility of ascribing experiences to an identical subject. Given this definition, it becomes tempting for the Kantian to abbreviate his answer to (a) by saying that what the idea of experience of objects requires is the unity of consciousness; in other words, given his argument concerning (a), and his definition of the unity of consciousness, it begins to appear that (a) and (b) have the same answer. But there is no reason to accept Kant's answer to (b). The question "what explains the unity of consciousness?" is the question "what unites the experiences had by a single subject?" This is not a question to be settled by definitional maneuvers. In terms of the dual perspective

approach, it is an *external* question, and there is no reason to suppose that the materials used in answering (a) must also be used in answering (b). Even if the thinking of self-ascriptive thoughts is the key to (a), the possibility of ascribing experiences to subjects need not be the key to (b). According to the Reductionist, the keys to (b) are the R-relatedness and possible co-consciousness of mental states. Consequently, the Reductionist should argue, he can accept Kant's answer to (a), while continuing to insist upon an explanation or description of a given mental life in purely impersonal terms.

What would be the consequences of this Reductionist response to Kant for the charge that the Reductionist position is covertly eliminative with respect to subjects of experience? This question points to a difficulty for the Reductionist, although the precise form of the difficulty will depend upon whether or not he regards his external, sideways on perspective as, in any sense, privileged. If it is privileged, and if the Reductionist does *not* regard the Kantian argument as calling for the *sideways on* ascription of experiences to subjects, conceived of as distinct from those experiences, then it is no clearer now than it was in section 1 how the Reductionist is to motivate the claim that the existence of subjects of experience is more than an illusion generated by the way in which we talk. So the Reductionist faces the following dilemma: either he accepts that the Kantian argument does require the ascription of experiences to subjects, even from the privileged, external perspective, in which case he must give up the claim that the unity of a mental life can be described and explained in impersonal terms. Alternatively, he can continue to insist that claiming that experiences are had by subjects is, despite Kant's argument, explanatorily redundant from the privileged perspective, in which case he is powerless to resist the charge that his commitment to the existence of subjects is less than securely grounded. Parfit's response to Descartes is a clear reflection of this dilemma. He describes the Cogito argument as true but misleading; however, the *only* sense in which what Descartes says is true is that "we do ascribe thoughts to thinkers"<sup>41</sup>—that is to say, this is how we talk. The sense in which what Descartes says is misleading is that it does not establish the need for mental states to be ascribed to subjects from the Reductionist's external stand-

---

<sup>41</sup> See, again, Parfit's discussion of Descartes in *Reasons and Persons*, 224–26.

point. But if the Reductionist's perspective is a privileged one, then it seems a short step to the claim that the very idea of there being subjects of experience is, strictly speaking, erroneous.

One response to this might be to argue that although the Reductionist's right to the claim that there are subjects of experience has yet to be established, this is not the same as saying that the Reductionist is committed to *denying* their existence. Perhaps he must remain neutral on the question, or suggest that there might be some other way of motivating the claim that the existence of subjects of experience is something more than a linguistic illusion. On reflection, however, it should be clear that this does not defuse the original dilemma. For the stronger the alternative grounds which the Reductionist provides for not viewing reference to persons as, from the privileged perspective, explanatorily redundant, the greater will be the pressure on the claim that we can give a complete description of reality without claiming that persons exist. Perhaps, with a sufficiently strict interpretation of the notion of completeness, there might be some way of reconciling (9) with an argument for the explanatory indispensability of sideways on references to persons, but it is unappealing to require the somewhat hazy notion of completeness to bear such weight.

An alternative response would be to deny the privileged status of the external standpoint. If the ascription of experiences to a subject is, from a standpoint within a given mental life, a vital part of what makes that mental life internally intelligible, then, to that extent, the existence of subjects to which experiences may be ascribed is as deep a truth as one could want it to be. It could not merely be a linguistically generated illusion. In that case, however, what would be the force of the claim that there is *also* a detached standpoint from which the unity of the life can be explained in entirely different terms? If the sideways on perspective is not privileged, and merely represents one among other ways of looking at things, *and* if it is not a way of looking at or thinking about a mental life which would render that life intelligible "from the inside"—at least if the life is to include experiences conceptualizable from the inside as experiences of spatial objects—then *why* should the detached description be of any *interest* to us?

It should not be assumed that these questions are unanswerable. Perhaps the Reductionist ought to regard his view as being motivated by its ethical consequences. In that case, however, it would

be the ethics which would underpin the theory of the self rather than vice versa, as might have been supposed. What should be conceded is that the Reductionist has more work to do if he is to repudiate the charge that he is an error theorist who is unwilling to come clean. At a certain level, the Reductionist's stance is clearly revisionary, but he wants to argue that he is able to accommodate the best insights of other views, including Kant's, without conceding any significant ground. It is this claim which remains very much open to question.<sup>42</sup>

*Wadham College, Oxford*

<sup>42</sup> For helpful comments, I am much indebted to John Campbell, David Charles, Naomi Eilan, Sir Peter Strawson, and Richard Swinburne.