CONTEMPORARY PHILOSOPHY

General Editor

Max Black, Cornell University

Editorial Committee

Charles A. Baylis, Duke University
William Frankena, University of Michigan
Morton White, Harvard University

Induction and Hypothesis: A Study of the Logic of Confirmation. By S. F. Barker.

Perceiving: A Philosophical Study. By Roderick M. Chisholm.

The Moral Point of View: A Rational Basis of Ethics. By Kurt Baier.

Religious Belief. By C. B. Martin.

Knowledge and Belief: An Introduction to the Logic of the Two Notions. By Jaakko Hintikka.

Reference and Generality: An Examination of Some Medieval and Modern Theories. By Peter Thomas Geach.

Self-Knowledge and Self-Identity. By Sydney Shoemaker.

SELF-KNOWLEDGE AND SELF-IDENTITY

By Sydney Shoemaker
CORNELL UNIVERSITY

Cornell University Press
ITHACA, NEW YORK

tempted to discuss the possibility of personal survival after death, or the relation of the notion of personal identity to the religious notions of personal immortality and bodily resurrection.

This book grew out of my doctoral dissertation, which was presented to Cornell University in 1958. A paper based on it was read at a symposium on self-identity at the meetings of the Eastern Division of the American Philosophical Association in 1959, and printed, together with Terence Penelhum's contribution to the symposium, in the *Journal of Philosophy*, LVI (October 22, 1959). Brief portions of that paper are incorporated into the present work, and I am grateful to the editors of the *Journal of Philosophy* for permission to reproduce them here. The present version was written mainly during 1960–1961, when I held the Santayana Fellowship at Harvard. I wish to express my thanks to Harvard University for this fellowship, to the Harvard Philosophy Department for their kind hospitality during my year in Cambridge, and to the Cornell Graduate School for a grant covering the cost of the final typing of the manuscript. The earliest versions of this work were read by Professor Norman Malcolm, to whom I am extremely grateful for his many valuable criticisms and for his constant encouragement. I am also grateful to other friends and colleagues who have read and criticized portions of the manuscript or discussed with me the ideas in it; special thanks are due Edmund Gettier, Carl Ginet, Norman Kretzmann, and Nelson Pike. Finally, I wish to thank Professor Max Black, editor of this series, for his encouragement and helpful advice. For whatever inadequacies and mistakes are contained in this work I am of course entirely responsible.

SYDNEY SHOEMAKER

Ithaca, New York
December 1962

Contents

unknown, foreign to me, than what is real, what I am aware of—my very self." [26]

Descartes goes on to argue, of course, that in fact we don't understand "corporeal objects" better than we understand our own nature. But his confession that he could not help thinking otherwise is revealing. Most of us have the feeling, which Descartes had but resisted, that we do not really understand the nature of a thing unless we can imagine or picture it. The ordinary way of picturing a person is by picturing the human face or the human body. And Wittgenstein remarked that "the human body is the best picture of the human soul." [27] But once one has the idea that a person is something logically distinct from his body, this way of picturing a person no longer seems to do. It is, among other things, the apparent lack of a philosophically suitable way of picturing a person that makes persons seem a mysterious sort of objects. One way in which a consideration of the nature of self-knowledge gives rise to problems about the nature of persons is by making persons seem mysterious in this way. As we shall see, there are others.

[26] "Meditations," p. 71.

[27] Ludwig Wittgenstein, *Philosophical Investigations* (Oxford, 1953), p. 178.

Two

Are Selves Substances?

1. Philosophical discussion of the nature of persons or selves has often centered on the question "Is the self a substance?" What appears to be the same question is sometimes expressed by asking whether the self is (or, to improve on the traditional formulation, whether *a* self is) a "pure ego" or a "subject." This question can be divided into two parts, the question whether at each point in a person's life history there is a substance to which his states at that time belong, and the question whether the identity of a self through time consists in the identity of a single substance. The term "subject" is most frequently used in formulations of the first of these questions, but it is sometimes used in formulations of the second, and I believe that in discussions of the self the expressions "substance," "pure ego," and "subject" can generally be regarded as synonymous.

It is not easy to make sense of the controversy to which this question has given rise. The statement "The self is a substance

(pure ego, subject)" can easily seem, and I think has been regarded by many of its defenders as being, an obvious truism. Yet its truth has often been disputed. Part of this chapter will be devoted to showing that no clear sense has been given to the denial that selves are substances. But it is not enough to show this; if we are to understand many of the questions that have been raised about the nature of the self, and many of the theories that have been advanced, we must show how the denial that selves are substances has *seemed* to acquire a sense, and how philosophers have been led to make it. To show this is the primary aim of this chapter.

2. In his "Second Meditation" Descartes concludes that he is not "that set of limbs called a human body" and goes on to ask what he *is*: "What then am I? A conscious being. What is that? A being that doubts, understands, asserts, is willing, is unwilling; further, that has sense and imagination." [1] Thomas Reid raises a similar question and gives a similar answer. Speaking of himself he says "Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers." [2] These remarks by Descartes and Reid exemplify what is surely the most natural way of expressing the conclusion of the line of argument presented in the preceding chapter. The considerations that lead us to the conclusion that a person is *not* a body seem to indicate, in a very general way, what a person *is*. We are led to the conclusion that a person is something distinct from a body by taking as paradigms of statements about persons, and statements expressing knowledge of persons, what I have called "first-person psychological statements." One might say that we know at least one thing about what a person is, namely that it is the sort of thing that sort of statement is about. Since typical first-person psychologi-

[1] "Meditations," p. 70.

cal statements are "I have a headache," "I am thinking about this book," and "I want to go home," it is natural to say, as Descartes or Reid would say, that what a person is, essentially, is something that experiences (has experiences), thinks (has thoughts), desires things (has desires), and so on. This can be expressed by saying that a person is something whose features (attributes, properties) are psychological features.

If one says that a person is *essentially* something that thinks, experiences, and so on, or that a person is something that thinks, experiences, and so on, *and* is distinct from a body, or that a person is something *all* of whose features are psychological features (and not physical features), one is certainly not asserting a mere truism. What one is asserting, whether true or not, is not *obviously* true, and is philosophically controversial. But it would seem that at least part of what one is saying *is* an obvious truism, namely that a person is *at least* something that thinks and experiences, something that has "psychological features." For what is this but to say that persons think, have pains, desire things, and so forth? What should be regarded as controversial, it would seem, is not the assertion that persons have psychological features, but the assertion that what has these features is something distinct from a body, something that does not have physical features as well.

Both Descartes and Reid say that persons (selves) are substances. But in saying this, I think, they did not regard themselves as adding anything to the assertion that a person is something, a "being," that doubts, thinks, understands, deliberates, and so forth. The claim that a person is a substance is sometimes expressed by saying that a person is a subject of thought and experience. It is difficult to see what can be meant by the expression "subject of thought and experience" if it does not mean "something that thinks and experiences (has experiences)." But if it means this, and if persons think

[2] *Essays*, p. 203.

and have experiences, it follows that a person is a subject of thought and experience, and therefore a substance. To say that a person is a substance in this sense is but another way of expressing the truism mentioned above. We might distinguish this assertion from the view of which I have said that it is controversial and not a truism, the view that seems to follow from the line of argument presented in the preceding chapter, by describing the latter, in traditional terminology, as the view that a person is a *mental* substance, a *spiritual* substance, or an *immaterial* substance.

Now the latter view has typically been regarded by philosophers as but one of several "theories" about the nature of the the self, and has been held by many philosophers to be only possibly true, or to be false, or to be absurd, or to be utterly unintelligible. The view that these philosophers have questioned, or denied outright, is certainly questionable. Oddly enough, however, the part of it that is questioned has commonly been just the view of which I have said that it seems to be an obvious truism. What is challenged is the view that a person is a substance, not the view that a person is something mental or immaterial. Critics of the substance theory have often agreed with their opponents in identifying the problem of personal identity with what Broad called the problem of "the unity of the mind." [3] Hume, the most famous of these critics, held that a person is "nothing but a bundle or collection of different perceptions," [4] and on Hume's view the perceptions that constitute me are *my* perceptions, i.e., what would normally be regarded as contents of my mind. The question arises whether these critics, in denying that a self is a substance, meant to be denying that a person is something

[3] C. D. Broad, *The Mind and Its Place in Nature* (London, 1925), pp. 556 ff.

[4] *Treatise*, p. 252.

that thinks, perceives, feels pain, and so on. Reid apparently thought that they did. He says: "I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers." [5] One can imagine Reid saying against Hume: "I am not perceptions; I am something that perceives, or something that *has* perceptions."

One is inclined to say that if there is a sense of "subject" and "substance" in which it is a truism to say that a person is a subject or substance, this merely shows that those who have denied that a person is a subject or substance must have been using these words in some other sense. Whether this is so, however, can only be decided by an examination of how the philosophers in question have used, and explained the meanings of, these terms.

3. The first important philosopher to question the view that a self is a substance was Locke. He did not question this view because he doubted the existence of "immaterial substances." Throughout his discussion of personal identity Locke implies that when a person thinks there is always a substance that does the thinking. And he thought it "probable" that a person's consciousness "is annexed to, and the affection of, one individual immaterial substance," i.e., that when one and the same person thinks on two different occasions it is one and the same substance that does the thinking. But his opinion seems to have been that if a person's consciousness is always "annexed to" one individual substance, this is so only as a matter of contingent fact, not as a matter of logical necessity. Personal identity, while it may be correlated with identity of substance, does not consist in this, and it "matters not at all," so far as the nature of personal identity is concerned, "whether it be the same identical substance, which always thinks in the

[5] *Essays*, p. 203.

same person." It is conceivable, Locke thought, that personal identity can be "preserved in the change of immaterial substance, or variety of immaterial substances." [6]

But what, in advancing this view, did Locke mean by "substance"? He remarks at one point that "doubts are raised as to whether we are the same thinking things, i.e., the same *substance*, or no," and his wording here would suggest that a thinking thing *is* a substance.[7] And Locke repeatedly speaks of immaterial substances as entities that think "in" persons. But if Locke allows that to be a thinking thing is to be a substance, he is clearly in difficulty, as Reid pointed out.[8] For Locke defines "person" as meaning "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking being." [9] If persons are thinking things, and thinking things are substances, then persons are substances. And if it follows from the definition of "person" that a person is a substance, it is surely self-contradictory to say that the identity of a person does not involve the identity of a substance.

A similar contradiction occurs in Russell's *The Problems of Philosophy*. The "real self," Russell says at one point, "is as hard to arrive at as the real table, and does not seem to have that absolute convincing certainty that belongs to particular experiences." [10] When I see a brown color, "what is quite certain is not 'I am seeing a brown colour,' but rather, 'a brown colour is being seen.'" The latter "involves something (or somebody) which (or who) sees the brown colour; but it does not of itself involve that more or less permanent person whom we call 'I.'" And it may be, "so far as immediate certainty goes," that "the something which sees the brown colour is

[6] *Essay*, I, 465, 450, 453. [7] *Ibid.*, p. 450.
[8] See Reid, *Essays*, pp. 212–213. [9] *Essay*, I, 448.
[10] Bertrand Russell, *The Problems of Philosophy* (London, 1950), p. 19.

quite momentary, and not the same as the something which has some different experience the next moment." Here Russell says that what we call "I" is a "more or less permanent person," not the something (which may exist only momentarily) that sees brown colors and has other experiences. Later on, however, Russell says: "We know the truth 'I am acquainted with this sense-datum.' It is hard to see how we could know this truth, or even understand what is meant by it, unless we were acquainted with something which we call 'I.'" [11] And he goes on to say that "it does not seem necessary to suppose that we are acquainted with a more or less permanent person, the same today as yesterday, but it does seem as though we must be acquainted with that thing, whatever its nature, which sees the sun and has acquaintance with sense-data." Here, as in the earlier passage, Russell distinguishes between the "more or less permanent person," with which he thinks we need not be acquainted, and something else, the entity which is acquainted with sense-data (including, presumably, the "brown colours" of the previous passage), with which he thinks we must be acquainted. But whereas he had previously said that the word "I" refers to the "more or less permanent person," here he says that it refers to this something else.

Though Russell does not use the term "substance," his position is clearly similar to Locke's. Both distinguish between persons, which can be said to exist or persist for relatively long periods of time (as long, presumably, as persons are said to live), and entities of another sort. To the latter, which Locke calls "immaterial substances" and Russell elsewhere calls "subjects," they ascribe mental states and activities. It is these that are said to think and experience, and it is presumably to these, therefore, that the word "I" in psychological statements refers (as Russell at one point says). It is held by both Locke

[11] *Ibid.*, p. 51.

and Russell that the identity of a person does not necessarily involve the identity of an entity of this other sort; as Russell said in a later work, "nothing is to be assumed as to the identity of the subjects of different experiences belonging to the same person." [12] From this it would seem to follow that persons cannot be identified with substances or subjects. The contradiction I have attributed to Locke consists of a conjunction of this view with the commonplace assertion that persons think (this being part of his definition of "person"), and that attributed to Russell consists of a conjunction of this view with the equally commonplace assertion that it is to persons that the word "I" refers.

That Locke and Russell both contradict themselves in expounding this view is not in itself sufficient grounds for rejecting the view as inherently self-contradictory. But, as I shall try to show, this Lockian view (as I shall call it) can escape the charge of inconsistency only at the cost of becoming unintelligible.

4. Many philosophers have held the Lockian view to be true or possibly true. C. D. Broad suggests, as a possible view concerning the nature of the unity of minds, that "there is a different Pure Ego for each different total state of the same mind, and that two successive total states are assigned to the same mind because of certain characteristic relations which they have to each other and which they do not have to other total states which would not be assigned to this mind." [13] This view he contrasts with that sort of "Pure Ego theory" which holds that the different "total states" of the same mind belong to one and the same pure ego, which persists as long

[12] Bertrand Russell, "On the Nature of Acquaintance," *Logic and Knowledge: Essays, 1901–1950*, ed. R. C. Marsh (London, 1956), p. 163.

[13] *Mind and Its Place*, p. 561.

as the mind persists. The same contrast is made by G. E. Moore. It *may* be, said Moore, that

> the relation which unites all those acts of direct apprehension which are mine, and which is what we mean to say they have to one another when we say that they are all mine, really does consist in the fact that one and the same entity is *what* directly apprehends in each of them: in which case this entity could properly be called "me," and it *would* be true to say that, when I see this black mark, *I* directly apprehend it.[14]

But it is also possible, he says, that

> the entity which directly apprehends, in those acts of direct apprehension which are mine, is numerically different in every different act; and that what I mean by calling all these different acts *mine* is either merely that they have some kind of relation to *one another* or that they all have a common relation to some other entity, external to them, which may or may not be something which deserves to be called "me."

Now what have these philosophers meant by the terms "substance," "subject," and "pure ego"? Moore, of course, did not use any of these terms; he simply uses the expression "entity which directly apprehends," and says that it is possible that the entity that directly apprehends one of my sense-data is not the same as the entity that directly apprehends another of my sense-data. But his "entity which directly apprehends" is pretty clearly the same as Russell's "subject," for Russell at one point defined "subject" as meaning "any entity which is acquainted with something." [15] Russell's subject is *that which* is acquainted with things (e.g., sees brown colors), just as Locke's substance is *that which* thinks. And this is *all* that we

[14] "The Status of Sense-Data," *Philosophical Studies* (London: Routledge & Kegan Paul, Ltd., 1922), pp. 174–75.

[15] "On the Nature of Acquaintance," p. 162.

are told about these entities. Broad defines "pure ego" as meaning "a particular existent which is of a different kind from any event; it owns events but is not itself an event." But in what sense does a pure ego "own" a mental event? All that Broad tells us about this "ownership" is that it is "a peculiar asymmetric relation," and this is to tell us practically nothing.[16] If Broad thought that his definition of "pure ego" was satisfactory, he must have thought that everyone is already familiar with the relevant sense of "own." And he can have thought this, I believe, only if he assumed that "own" would be taken as a synonym of "have," and was thinking of the use of "have" in such sentences as "I have a toothache" and "He just had an idea." If this is so, a pure ego, as Broad defines the term, is simply something that has thoughts, feelings, desires, and so on.

Thus it would appear that we are to understand the terms "substance," "subject," and "pure ego" as meaning: whatever is designated by the grammatical subject of a psychological statement, i.e., anything to which psychological attributes belong, e.g., anything that thinks, has experiences, or directly apprehends (is acquainted with) objects. That this is the intended meaning of these terms is indicated also by the sorts of things that have been counted as grounds for thinking that there are such entities. Broad classifies pure ego theories as "central theories" concerning the unity of the mind, and says:

> The *prima facie* presumption in favor of Central theories and against Non-Central theories is the common usage of language, which strongly suggests the existence of a Centre. We say: "I am thinking of this book, and wanting my tea, and feeling tired, and remembering the tie that my friend wore yesterday." This certainly suggests that "I" is the proper name of a certain existent

[16] *Mind and Its Place*, pp. 558, 562.

> which stands in a common asymmetric relation to all those contemporary mental events.[17]

Russell, after he had given up the view that there are subjects, gave a similar account of why it is thought (mistakenly, according to him) that there are such things. "We say: '*I* think so-and-so,' and this word 'I' suggests that thinking is the act of a person." [18] What gives rise to the belief in subjects or pure egos, according to Broad and Russell, is the use of the word "I" as a grammatical subject in psychological statements, which "suggests" that there is something to which this word refers, something that is the subject of mental acts and states.

5. The accounts given by Broad and Russell, as to why it is thought that there are subjects or pure egos, seem essentially correct. There is surely a sense in which Reid is appealing to "the common usage of language" when he says "I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers." [19] Reid apparently regards the statement "I am something that thinks" as following necessarily from those statements, like "I am thinking of this book," which according to Broad and Russell "suggest" its truth.

But Broad and Russell, unlike Reid, are among those who have held that the subject of thought, if there is such a thing, may have only a momentary existence and may not be identical with a "more or less permanent person." And their account of the grounds for the belief in subjects renders this view incoherent. If, as Russell says, the "I" in "I think so-and-so" suggests that thinking is the act of a person, it surely does not suggest that thinking is the act of an entity which is not, or may not be, a person. Broad says that the statement "I

[17] *Ibid.*, p. 584. [18] *Analysis of Mind*, p. 18. [19] *Essays*, p. 203.

am thinking of this book, and wanting my tea, and feeling tired, and remembering the tie that my friend wore yesterday" suggests that "I" is "the proper name of a certain existent which stands in a common asymmetric relation to all those contemporary mental events." But if this is so, the statement "I am thinking about this book now and was thinking about it yesterday" must surely suggest that "I" is the name of a *persisting* existent, one which stands, at different times, in a "common asymmetric relation" to noncontemporary mental events. Suppose that Jones now says "I thought about this book yesterday," and that, as a matter of fact, Jones did think about the book yesterday. Surely the "common usage of language" dictates that in this case Jones's statement would be true. But if the "I" in Jones's statement were the name of a momentarily existing entity, something that did not exist yesterday and hence cannot be identified with the "permanent person" Jones, then his statement would be false.

The reference of the word "I" poses a serious difficulty for anyone who wishes to defend the plausibility of the Lockian view. If one is to be able to appeal to the common usage of language in support of the view that there is a subject of thought and experience, one must hold that the word "I," indeed, any expression that occurs as the subject in a psychological statement, refers to a subject. But one cannot hold this if one thinks that subjects are not, or may not be, persons, for it seems clear that the word "I" does refer to a person. Thus we find inconsistencies like that in Russell's *Problems of Philosophy*.

Let me now put in a different way the difficulties that seem to me inherent in any version of the Lockian view. The reasons that have been given for holding that there are "subjects," "pure egos," or "immaterial substances," as well as the definitions that have been given for these terms, indicate that

these expressions are intended to refer to entities corresponding to the grammatical subjects of psychological statements. Taking the word "subject," then, let us assume that it can be defined as meaning "something that has psychological attributes, e.g., something that thinks and feels." It would seem, offhand, that if we understand the expressions "think" and "feel" we should have no difficulty in understanding such a definition. And of course we do understand these expressions as they are ordinarily used. The difficulty is that, as they are ordinarily used, what can be said to think or feel is a *person*. Now, anyone who holds that there are subjects and that subjects cannot be identified with persons must be holding one of the following positions: (1) Strictly speaking, persons do not think or feel at all; only a subject can be said to think or feel. (2) Both subjects and persons think and feel, and the sense of "think" and "feel" in which persons think and feel is the same as that in which subjects think and feel. (3) There is one sense of "think" and "feel" in which persons think and feel, and there is a different sense of these words in which subjects think and feel. Let us consider these positions in order.

It is clear that (1) is altogether untenable. To hold it is to hold that any statement that says of a person that he thinks or feels something is senseless or necessarily false. But we all make such statements, and it is senseless to say that everyone misuses the words "think" and "feel"; if everyone is taught to use an expression in a certain way, and everyone does use it in that way, then that is *a* correct way of using it. The ordinary sense of "think" must be one in which a person can be said to think. Perhaps it will be said that while a statement like "That person is thinking" may be true, what it *really means* is not that a person is thinking but that a subject, which is in some way related to a certain person, is thinking. But

unless the words are being used figuratively (as normally they are not) such statements cannot fail to mean that a person is thinking; they mean what they say. Of course, it might be said that any such statement can be analyzed into a statement in which a subject is said to think in a sense of "think" in which persons cannot be said to think. But this is to assert position (3), not position (1).

According to (2), subjects think and feel, and persons think and feel, and both think and feel in exactly the same sense of the words "think" and "feel." On this view a person would necessarily be a subject, and it would be a flat contradiction to say that a person may persist when no subject persists. So if one holds (2), and holds also that every thought and experience must have a subject that is not a person, one can avoid self-contradiction only at the cost of holding that every thought and experience must have at least two subjects, both of them subjects in exactly the same sense of the word "subject"; each thought and experience will have one subject that is a person and another that is not a person, and it will belong to both of these subjects in exactly the same way. I am confident that nobody would want to hold this.

So we are left with (3). About this I shall make only two remarks. First, if one holds this position one cannot cite the fact that we all understand the expressions "think" and "feel" as grounds for saying that the definition of "subject" we are considering is perfectly intelligible. For according to (3), the familiar sense of "think," the sense in which persons can be said to think, is not the sense in which subjects think. Anyone who holds (3) owes us an explanation of the sense in which he is using these terms as applied to subjects, and to my knowledge no such explanation has ever been given. Because of the way in which philosophers have supported the claim that there is a subject of thought and experience, I think that they have

taken it for granted that the sense in which they have used terms like "think" and "feel" is the ordinary sense and therefore requires no explanation. But, and this is my second point, if one holds (3) one cannot support the claim that there are subjects in the way in which this claim has most commonly been supported. One cannot, that is, appeal to what Broad calls "the common usage of language." A statement like "I am thinking of this book" may suggest, or even entail, that there is something that thinks *in the ordinary sense of "thinks"*; it certainly does not suggest, let alone entail, that there is something that thinks in a sense of "thinks" that is not the ordinary sense and not the sense in which "thinking" is being used in that sentence.

6. The Lockian view, while it denies that a person *is* a subject (substance, pure ego), i.e., that the identity of a person necessarily involves the identity of a subject or substance, does not deny that there are subjects of thought and experience. I turn now to a consideration of a more radical view, one that makes just that denial. Those who have held this view have commonly held one version or another of what Broad has termed the "bundle theory" of the self, a theory that takes its name from Hume's famous remark that a person is "nothing but a bundle or collection of different perceptions." [20] The existence of a "perception," Hume apparently held, does not involve the existence of anything that perceives or has the perception. Pains and images are examples of what Hume calls perceptions, and he appears to have held that to say that a particular person has a pain is not to say that a pain belongs to a certain subject or substance, but is only to say that a pain is included in a certain "bundle," or is a member of a certain "collection," of perceptions.

[20] *Treatise*, p. 252.

The bundle theory of the self has also been called the "serial theory," the "associationalist theory," and the "logical construction theory," and has been widely held among empiricist philosophers. Russell held a version of it at one time. In his essay "On Propositions" Russell wrote that "the theory which analyses a presentation into act and object no longer satisfies me. The act, or subject, is schematically convenient, but not empirically discoverable."[21] This is developed at greater length in *The Analysis of Mind:*

> Empirically, I cannot discover anything corresponding to the supposed act; and theoretically I cannot see that it is indispensable. We say: "*I* think so-and-so," and this word "I" suggests that thinking is the act of a person. Meinong's "act" is the ghost of the subject, or what was once the full-blooded soul. It is supposed that thoughts cannot just come and go, but need a person to think them. Now, of course it is true that thoughts can be collected into bundles, so that one bundle is my thoughts, another is your thoughts, and a third is the thoughts of Mr. Jones. But I think that the person is not an ingredient in the single thought: he is rather constituted by relations of the thoughts to each other and to the body. . . . The grammatical forms "I think," "you think," and "Mr. Jones thinks," are misleading if regarded as indicating an analysis of a single thought. It would be better to say "it thinks in me," like "it rains here"; or better still, "there is a thought in me." This is simply on the ground that what Meinong calls the act in thinking is not empirically discoverable, or logically deducible from what we can observe.[22]

The implication here is that a person, or at least the mind of a person, is simply a "bundle" of thoughts (and, presumably, other mental events or objects). In "The Philosophy of

[21] *Logic and Knowledge*, p. 305.

[22] London, 1921, pp. 17–18. Reprinted with permission of George Allen & Unwin, Ltd., and the Macmillan Company of New York.

Logical Atomism" Russell says that a person is "a certain series of experiences."[23]

My arguments in Section 5 against the Lockian view might be summarized by saying that the terms "subject," "substance," and "pure ego," as used in discussions of the self, are either synonyms of the word "person" or without meaning, and that it is therefore either self-contradictory or meaningless to say that the identity of a person does not involve the identity of a subject. If that argument is correct, it would appear that the denial that there are subjects (substances, pure egos) is either unintelligible or else amounts to a denial that there are persons. But the bundle theory is commonly put forward, not as a denial that there are persons, but as a theory concerning the nature of persons. It would seem, then, that this theory, if it has any meaning at all, implies both that there are persons (for it tries to say what persons really are) and that there are not (for it denies the existence of subjects), and is therefore incoherent. Before accepting this verdict, however, we must take a closer look at what bundle theorists have said. For there is what seems at first sight a way of making the theory intelligible and consistent.

7. Russell said that "the grammatical forms 'I think,' 'you think,' and 'Mr. Jones thinks,' are misleading if regarded as indicating an analysis of a single thought." In the same passage he says that the fact that we say "I think so and so" suggests that thinking is the act of a person, and he implies that what this suggests is false. But it appears that he does not mean that whenever one makes a statement of the form "I think so and so" one is saying something false. It is one thing, apparently, for a sentence, or the grammatical form of a sentence, to be "misleading," for it to "suggest" something that is false,

[23] *Logic and Knowledge*, p. 277.

and another for the sentence itself (or the statement expressed by it) to be false. Russell seems willing to admit that there is a sense in which thoughts can be said to "belong" to persons, for he tries to say what this belonging consists in: "Thoughts can be collected into bundles, so that one bundle is my thoughts, another is your thoughts, and a third is the thoughts of Mr. Jones." But he thinks that the form of our psychological statements (first person and third person alike) misleads us by suggesting that thoughts belong to persons in some *other* sense.

It is not clear, however, what is meant by saying that the form of a sentence suggests something other than what that sentence says, nor is it clear what, according to Russell, the grammatical form of "I think" suggests. It obviously does not help to say that what is falsely suggested is the existence of a subject of thought and experience, for what is in question is precisely the intelligibility of saying that this "suggestion" is false. One thing that Russell believes to be suggested by such grammatical forms, and is concerned to deny, is that thinking is an "act." But this gets us no farther. What is it that thinking is not? With what is it being contrasted when it is denied that it is an act? The grammatical form "I think" could hardly be said to suggest that thinking is a physical act, like kicking or hitting; what Russell must be denying is that thinking is a mental act. But he neglects to tell us what it is (or would be) for something to be a mental act. If we were to compare the verb "think" with other psychological verbs, we might find reasons for saying that it, unlike some of the others, is not used to report the occurrence of discrete acts. If, for example, we take *saying something to oneself* as our paradigm of a mental act, we will doubtless conclude that *thinking that* so and so is the case (i.e., believing) is not a mental act, and we may also conclude that *thinking about* something is not an

act and cannot be resolved into a series of acts. But Russell was not thinking along such lines as these. It is clear that he held that there are no mental acts at all, and not merely that there is no mental act of thinking. He could not have explained the sense in which he was using the word "act" by giving us an example of what he would count as a mental act. What seems likely is that he denied that thinking is an act because he supposed, reasonably enough, that an act requires an agent, and because he wanted to deny the existence of an agent in thought, i.e., of something that thinks. But, again, it is precisely the intelligibility of the latter denial that is in question, especially since Russell apparently does not wish to deny that statements like "Jones is thinking" are often true.

But I have been ignoring an important phrase in the passage quoted from Russell. He says that the grammatical form "I think" is misleading *if regarded as indicating an analysis of a single thought*. Perhaps his view—and that of bundle theorists in general—is really a thesis about how psychological statements are to be *analyzed*, and perhaps the denial that there is a subject of thought and experience is not essential to it. Such, at first sight, seems to be the nature of the view advanced by A. J. Ayer in *Language Truth and Logic:*

> We do not deny, indeed, that a given sense-content can legitimately be said to be experienced by a particular subject; but we shall see that this relation of being experienced by a particular subject is to be analysed in terms of the relationship of sense-contents to one another, and not in terms of a substantival ego and its mysterious acts.[24]

Ayer seems to be saying that it is the *meaning* of the statement "There is a subject of experience," not its *truth*, that is at issue. This is encouraging, for we have seen that the truth of

[24] London, 1946, p. 122.

that statement seems beyond question. But what, then, are we to make of Ayer's phrase "not in terms of a substantival ego and its mysterious acts"? What does he mean by "substantival ego"? Ayer also speaks of "the substance which is supposed to perform the so-called act of sensing," saying that it is impossible to verify the existence of such an entity, and it is pretty clear that he uses that phrase to mean the same as "substantival ego."[25] So Ayer seems to be saying that while there are *subjects* which experience sense-contents, it is false (or senseless to suppose) that there are *substances* which sense these sense-contents. And this sounds, on the face of it, like a self-contradiction.

But we must pursue further the suggestion that the bundle theory, with its seemingly paradoxical denial that selves are substances, is really a theory concerning the correct analysis of statements about persons or selves. A sophisticated bundle theorist (or logical construction theorist) might state the matter as follows: There is a sense of "A self is a substance" in which what it expresses is an incontrovertible truism. For this sentence may mean simply that statements like "I am thinking about this book," "Jones has a backache," and "The person who has a severe headache now is the same as the person who had a pain in his foot yesterday" are significant and sometimes true. When it means this, the bundle theorist has (or should have) no quarrel with it. But the sentence "A self is a substance" has another possible meaning. It might be used to mean that such psychological statements, in addition to being significant and sometimes true, are, in the forms in which they are ordinarily expressed, unanalyzable (or "fully analyzed"). This the bundle theorist does deny, and this he can deny without being open to the charge of having denied an obvious truism. The bundle theorist holds that the correct analysis of

[25] *Ibid.*

a proposition about a person will always result in a sentence radically different in form from the sentences by which the proposition would ordinarily be expressed. In denying that a self is a *persisting* substance he is maintaining that the final analysis of a proposition that would ordinarily be expressed in a sentence of the form "The person who had Φ at t_1 is the same as the person who had Ψ at t_2" will always result in a sentence that does not have the form of an identity statement (here he is contradicting the view of those substance theorists, like Reid and Butler, who have held that personal identity is "indefinable"). And in denying that there is a subject of experience he is maintaining that the final analysis of a proposition that would ordinarily be expressed in the form "A is Φ" (where "A" is an expression referring to a person, and "Φ" is a psychological predicate), will always turn out to be a sentence in which expressions that would ordinarily be said to refer to persons ("I," "Jones," "the person who . . . ," and so on) do not occur, and one for whose subject expressions and individual variables these "person-referring expressions" could not significantly be substituted. The analysis of "I see an image," for example, would yield a sentence that does not contain "I" or any other person-referring expression, and does not contain any verb or predicate which, like "see," requires a person-referring expression as its grammatical subject.[26]

This account, of course, can be no clearer than the notion of analysis and the distinction between statements that are "analyzable" and statements that are "fully analyzed." In fact, I believe, this notion is far from clear, but I shall not argue the point here. It is worth noting, however, that some explanations given of this notion make the account given above circular, namely those that invoke a correspondence

[26] See H. P. Grice, "Personal Identity," *Mind*, L (1941), 334, where essentially this account is given.

theory of meaning and explain the phrase "fully analyzed sentence" by saying that a fully analyzed sentence is one whose elements ("simple symbols") stand for "elements of reality." For what then turns out to be meant by the claim that ordinary psychological sentences are analyzable is that person-referring expressions do not stand for, are not "logically proper names of," real particulars. But to say that person-referring expressions do not stand for real particulars, i.e., that persons are not real particulars, seems to be just another way of expressing the very claim whose meaning is to be explained, the claim that persons are not subjects or substances, and seems no less paradoxical than the original claim. The statement "Persons are particulars" seems at least as much a truism as the statement "Selves are substances." And if we try to distinguish two senses of the former sentence (as was done earlier with the latter one), and try to specify a nontruistic sense of it by using the notion of analysis, then we shall find ourselves either moving in a circle or involved in an infinite regress—unless, of course, we introduce another way of explaining the notion of analysis.

But whatever (if anything) turns out to be meant by the claim that psychological statements are analyzable, it is clear that those who have denied that selves are substances, or that there is a subject of thought and experience, did not first become persuaded that psychological statements are analyzable and then assert, on the basis of this, that selves are not substances. Rather, I think, they first became convinced that there is no subject (are no substances), and then tried to defend the intelligibility of this view by interpreting it as the view that psychological statements are analyzable. What these philosophers typically say is not "These statements are analyzable, for here is an analysis of one," but rather "There is no subject for the word 'I' to refer to, so these sentences (since they are obviously significant) *must* be analyzable (though I don't

know just what the analysis of any of them is)." What needs to be explained is how, prior to its interpretation in terms of the notion of analysis, the denial that selves are substances has seemed to philosophers to make sense and why many philosophers have thought it to be true. This I shall try to explain in the following sections. I shall try to show that our language, together with certain philosophical conceptions, seems to force on us certain pictures as representations of psychological facts and facts about the identity of persons, and that if one reflects on these pictures, and on the question of how persons can have knowledge of themselves, one can easily be led to deny that selves are substances.

8. It is a common philosophical view that every contingent fact a person knows he either knows directly on the basis of what he is presently observing or remembers having observed in the past, or else knows "inferentially" on the basis of his present and past observations. And it has generally been supposed that what I have called "first-person psychological statements" are such that a person who makes such a statement can be said to know it to be true. It is supposed, indeed, that when a person makes such a statement, and is not lying, he knows it to be true with the highest degree of certainty. But these statements are certainly contingent; it is contingently true that I now have a backache, that I am thinking about this book, and that I wonder when it will stop snowing. It is therefore plausible to suppose that if a person knows such a statement to be true he knows this in one of the ways just mentioned. But for anything to be known in any of these ways there must be *some* things that can be known solely on the basis of what the knower observes at the time at which he knows. And it is natural to suppose that the sort of fact that can be known in this way is precisely the sort of fact that is

expressed in a first-person psychological statement. My knowledge that I have a backache seems absolutely certain, and it certainly does not seem that in saying that I have a backache I am relying on my memory or making any sort of inference. But now it appears that when one knows the truth of a psychological statement about oneself, one must be in some sense observing or perceiving something. One need not, of course, be observing any material object. But involved in the view I am describing is the idea that *some* sort of perception, or something *like* perception, occurs whenever a person says (truthfully) that he thinks something, or feels something, or wants something. To this sort of perception philosophers have given a variety of names. They have called it "awareness," "consciousness," "immediate apprehension," "direct perception," and "acquaintance." [27]

If knowing that I am in pain involves some sort of perception or observation, it must involve that I perceive or observe *something*. Perception requires an *object*. There seems, offhand, to be no difficulty here. It is natural to say that when I know I am in pain what I perceive is the *pain*, and that when I know I have an image what I perceive is the *image*. There is not an equally natural answer to the question "What do I perceive when I know that I am thinking about this book?" but if this is also a case of empirical knowledge, as it is plausible to regard it, then it seems that here too there must be something that I perceive.

But it is also the case that nothing can be perceived or observed unless something observes or perceives it. Perception requires a *subject* as well as an object. So if my knowledge that I am in pain is based on some sort of observation or perception of a pain, there must be something that observes or perceives the pain. And since it is *I* that knows that I am in pain, it must be *I* that observes or perceives the pain.

[27] Since I think the notion that philosophers intend to express by these terms is modeled upon the notion of perception or observation, I shall allow myself, in this and the two following chapters, to use the terms "perceive" and "observe" to express it. Eventually I shall argue, in Chapter Six, that it is senseless to say that we perceive or observe our pains, images, and the like, and that our so-called "awareness" of, or "acquaintance with," mental contents cannot be regarded as like observation at all and cannot be said to explain our knowledge of first-person psychological statements.

Russell defined the term "subject" as meaning "any entity which is acquainted with something." Now persons are said to be subjects of other things besides acquaintance (or what I am calling perception); for instance, they are said to be subjects of thought and desire. But I want to suggest that the philosophical notion of a person as a subject has often been the notion of a person as primarily the subject of *perception* (acquaintance, awareness) and as only secondarily the subject of other things. This notion goes naturally with the views about knowledge described above. I am the subject of desire, for example, only when I desire something, which is not always. But on the view being considered, no matter what first-person statement I make, whether it be "I have a headache," "I want to sleep," or "I think it will rain," I must be observing or perceiving something. This is so, on that view, because when I make such a statement I know it to be true, and because I can only know it to be true on the basis of some sort of perception or awareness. Hume, it may be noted, remarks that "To hate, to love, to think, to feel, to see; all this is nothing but to perceive." [28]

In representing pictorially what is expressed by a statement of the form "S perceives O," it is natural to use a diagram like that in Figure 1. The concept of perception lends itself to

Figure 1

[28] *Treatise*, p. 67.

pictorial representation in a way in which other psychological concepts do not. If one were asked to draw a picture of a man thinking about philosophy, one might draw any number of different things; one might draw a man with a furrowed brow, a man scratching his head, a man staring at the floor, and so on. If asked to draw a picture of a man having a severe pain one might draw a man grimacing or a man with his body "doubled up." These pictures would not reflect the grammar of the sentences used to describe the states of affairs they represent (the sentences "He is thinking about philosophy" and "He has a severe pain"); there would not be in any of them a distinct thing representing what is designated by the grammatical object of the corresponding sentence, and there would therefore be nothing in the formal structure of the pictures to indicate the relational or quasi-relational character of these sentences. But if one were to draw a picture of a man perceiving something, e.g., a man seeing a tree, one would be likely to produce a picture similar in form to Figure 1. The picture would probably contain both a picture-man and a picture-tree; the grammatical subject and the grammatical object of the sentence "He sees a tree" would be represented in the picture by distinct things, and the relation of seeing would be represented by the spatial relationship between the two.

Of course, the objects of the sort of "perception" I have been talking about (which is essentially Russell's "acquaintance") are not material objects like trees; they are "mental objects" like images and pains. Nor is the subject what would be represented as subject in the picture of a man seeing a tree, namely a human body or a pair of human eyes. Nevertheless, I think that our tendency to represent the perception of these "mental objects" as in Figure 1 is to be accounted for by the fact that we take as our paradigm of perception the case of visual perception, or seeing, and that the sort of "seeing" we are likely to be thinking of is not, initially, the seeing of such

things as afterimages, but is rather the seeing of material objects. We might say that the picture we use to represent the perception (awareness) of mental objects is modeled upon the sort of picture we would draw if asked to draw a man seeing a material object.

On the conception I am discussing, a person is a subject, and a subject is essentially a perceiver of various kinds of mental objects. All the images, sensations, thoughts, and the like that a person is said to "have" at a given time are, according to this conception, related to a common subject by what Russell calls "acquaintance" and what I have been calling simply "perception." This may be represented as in Figure 2. The *S* represents the subject, and each O represents some mental object,

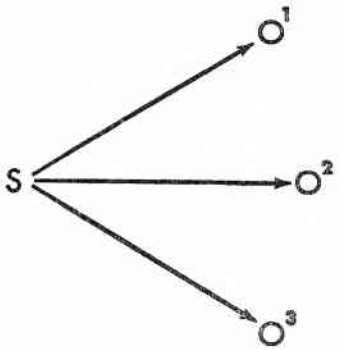

Figure 2

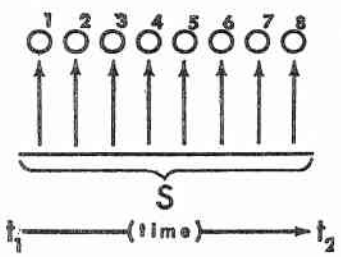

Figure 3

e.g., a pain, thought, or image, of which the person is aware at a given time. The arrows represent the relation of acquaintance or perception by which the subject is related to the objects. But since it is part of this conception that the subject *is* the person, the same subject must persist throughout the person's history. This can be represented as in Figure 3. The persistence of a single subject from t_1 to t_2 is represented by an unbroken line, and the fact that at different times during

that interval a number of different objects are perceived by that subject is represented by the fact that different *O*'s are connected to that line by arrows.

I shall now try to show how such pictures as Figure 2 and Figure 3 can seem to give sense to, and grounds for, the doubts that philosophers have raised concerning the assertion that a person is a subject. I shall consider first the Lockian doubt whether the same subject persists throughout the history of a person.

9. Figure 3 is supposed to represent the history, we might say the mental history, of a person. But does it represent the history of a person as that person himself knows it, i.e., as he remembers it? Consider the following analogy. Imagine that we have before us a number of glass beads strung on cords, and that we wish to determine whether two of the beads are strung on one and the same cord. Since both the beads and the cords are visible, we can start with one of our beads, follow along the cord on which it is strung, and see if we come eventually to the other bead. But what if our vision is obstructed and part of a strand of beads is hidden from us? Then we may see bead A, and the section of cord on which it is strung, and bead B, and the section of cord on which it is strung, and yet be unable to see whether A and B are strung on one and the same cord. Now let us compare seeing a bead and the section of cord on which it is strung with remembering a mental object, say an image, and the subject to which it belonged. To say that two different beads are on one and the same cord will correspond to saying that two different objects (say an image and a pain), occurring at different times, belong to one and the same subject. What will then correspond to having an unobstructed view of the strand of beads, presumably, will be having an uninterrupted memory, i.e., being able to remember every moment of a certain interval of time.

If a person had such a memory of his past history, it would seem that we could use Figure 3 to represent, not merely his past history, but the past history that he remembers. And then, it appears, there would be no difficulty about his knowing of two remembered mental objects that they belonged to one and the same subject; it would be like the case where one has an unobstructed view of a strand of beads, and can thus see that two different beads are on one and the same cord. But in fact our memories are full of gaps. One does not remember the intervals in one's past during which one was asleep, and much of one's past is simply forgotten. So the history a person actually remembers cannot be represented by Figure 3; it must be represented by something like Figure 4. But now one seems to be in a position, with regard to one's own past history, of the person who has an obstructed view of a strand of beads. It may be that S_1, S_2, S_3, and S_4 are all one subject (or segments of the history of one subject), but how can one know that they are if one has no memory of the intervals indicated by the gaps? One's memory, it appears, cannot tell one (except in the rare case in which one has an uninterrupted memory) whether a pain occurring at one time and an image occurring at another, e.g., a pain one remembers and an image one sees, belong to one and the same subject. And now one seems to be faced with the alternative of either admitting that one normally has no good grounds for ascribing any past mental object to oneself, even when one has a clear memory of it, or else denying that in order to know that a past mental object (for instance, a pain) was one's own one must know that it's subject is the same as the subject of one's present mental objects.

In his general discussion of the notion of substance Locke speaks as if substances are in principle unobservable. It would therefore seem that for Locke even Figure 4 cannot represent the history that a person remembers. Yet Locke sometimes

seems to argue along the lines indicated in the paragraph above. He says that no one would have had any reason to doubt that it is always the same substance that thinks in a person if our "perceptions, with their consciousness, always

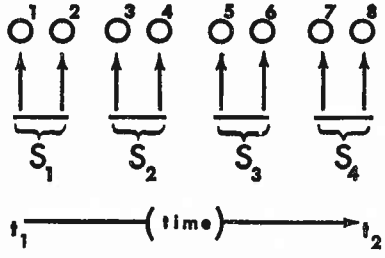

Figure 4

remained present in the mind, whereby the same thinking being would always be consciously present, and, as would be thought, evidently the same to itself." [29] But in fact, he says,

Figure 5

we never have "the whole train of all our past actions before our eyes in one view." Therefore, "our consciousness being interrupted, and we *losing sight of our past selves,* doubts are raised whether we are the same thinking being, i.e., the same

[29] *Essay,* I, 450.

substance, or no." [30] But the fact that our consciousness is interrupted, Locke thinks, is no reason for supposing that one cannot be sure that one is the same as the *person* one remembers doing a certain action, or having a certain thought, in the past. The very fact that one has the "same consciousness" of a past action as one has of one's present actions, i.e., remembers doing it, assures one that one is the person who did it. But since one can know with certainty that one is the person who did a certain action (or had a certain thought or experience), and cannot know with certainty that the substance that was the subject of that past action (thought, experience) is the subject of one's present actions and thoughts, it cannot be the case that one bases statements about one's own identity on facts about the identity of a substance, and it cannot be the case that the identity of a person *is* (consists in) the identity of a substance.

Locke's language is revealing. He speaks of our "losing sight of our past selves," and this suggests that he was thinking in terms of a visual analogy, such as my string-of-beads example or the diagrams in Figure 3 and Figure 4. But it is likely that he was also thinking in terms of an analogy of the following sort. Suppose that I have been given the job of following a certain person and recording his actions. And suppose first that I have followed him all day long and have never "lost sight" of him (have never "let him out of my sight," or "taken my eyes off him"). If now, at the end of the day, I remember having followed him all day, never letting him out of my sight, I can be sure that the person I see now is the same as the person I started following this morning. But suppose that just once during the day I let him out of my sight. Then the most that I can remember now is that I followed *a* person of a certain description until I lost sight of him, and that later on (perhaps only a minute or two later) I found *a* person of

[30] *Ibid.,* first italics mine.

the same description and have been following him ever since. That I remember all of this does not exclude the possibility that I have followed two different persons; it just might be the case that the man I see now is not the man I started following this morning, but is instead his identical twin. Similarly, I think that Locke is arguing that the most that I can remember, normally, is that *a* substance was the subject of certain actions or perceptions, and that later on *a* substance, perhaps exactly *like* the first one (if this has any meaning) was the subject of other actions or perceptions. If my memory is interrupted, or if I was not conscious throughout the period in question, there is nothing in what I remember that tells me whether or not these were one and the same substance.

All of this implies, of course, that if my memory were *not* interrupted then if one mental substance had been replaced by another just like it I would know that this had occurred; i.e., I would have detected the change when it occurred and would remember it now. Locke does not tell us what it would be like to observe, or remember, such a substitution of one substance for another. And as we saw in Section 5, if one inquires into the meaning of the terms "substance" and "subject," the idea that the subject of my present actions, thoughts, and so on may not be the subject of my actions and thoughts of yesterday seems either self-contradictory or unintelligible. All that I am saying now is that if one thinks in terms of certain pictures and analogies, which seem initially to be appropriate ones, this view does not seem at all absurd.

10. If, as Locke says, we frequently "lose sight" of our past selves, then at least we have something to lose. If the past history that a person remembers is represented by Figure 4 rather than Figure 3, then at any rate what a person perceives (is aware of) at a given time is *sometimes* what is represented

in Figure 2, for surely one cannot remember a subject existing at *t* if one was not aware of that subject at *t*.

But *can* Figure 2 represent what a person is aware of at a given time? What a person perceives are only the *objects* of perception (a tautology), so apparently the picture that represents what one perceives at a given time is not Figure 2 but Figure 5. But in that case *no one* ever perceives what is represented in Figure 2. If the person is myself, I do not perceive what corresponds to the *S*, and certainly no one else ever perceives a mental subject that is perceiving *my* pains, images, and thoughts. Yet, in a sense, Figure 2 suggests that the subject ought to be perceivable. For we perceive it *in the picture*. Figure 2 is an elaboration of Figure 1, and I have suggested that Figure 1 is modeled on the sort of picture we would draw if asked to depict a person seeing a material object. In pictures of the latter sort, however, what is represented as the subject of perception is in itself as capable of being perceived (though not by itself) as the object of perception. So we feel that the subject in Figure 1 and Figure 2 ought likewise to be perceivable by someone, and that if it is never perceived it must have the status of mermaids and unicorns, i.e., of things whose existence is doubtful because they are never perceived.

Hume, in a well-known passage, reports: "For my part, when I enter most intimately into what I call *myself* I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception."[81] In *The Problems of Philosophy* Russell held that we are sometimes acquainted with the subject of our own experience. Yet he remarks that this, while he thinks that it must be the case, does not *seem* to be the case, for "When we try to look into ourselves we

[81] *Treatise*, p. 252.

always seem to come upon some particular thought or feeling, and not upon the 'I' which has the thought or feeling." [32] Later on, as we have seen, Russell gave up the view that there are subjects on the grounds that the subject "is not empirically discoverable." [33] According to Ayer, the substantival ego is "an entirely unobservable entity" and "not revealed in self-consciousness." [34] And G. E. Moore remarks, in "The Refutation of Idealism," that "when we try to introspect the sensation of blue, all that we can see is the blue: the other element is as if it were diaphanous." [35] Moore is talking about the act of apprehension or sensation, not the subject of this act, but I think that he would say the same thing of the subject.

The remarks quoted above, with the possible exception of Ayer's, are offered as introspective reports. These philosophers claim that they have looked for something and failed to find it, and according to some of them this failure is grounds for saying that there are no entities of the sort they claim to have been looking for. But these claims must be treated with skepticism. If I am seen looking under tables and rummaging through wastepaper baskets, it will naturally be supposed that I am looking for something. But unless I have some conception of what the thing I claim to be looking for would look like, or of what it would be like to find it, my "looking" is mere show; whatever I may think I am doing, I am not looking for something. Likewise, a philosopher may close his eyes, furrow his brow, and "attend" ever so closely to the contents of his mind, but he cannot be said to be looking for the "I" (the subject of his experience) unless he knows what it would be like to find it, i.e., could identify a subject as such if he found one. None of the philosophers I have quoted, however, offers any account

[32] Page 50. [33] "On Propositions," p. 305.
[34] *Language Truth and Logic*, p. 126.
[35] *Philosophical Studies* (London, 1951), p. 25.

of what it would be like to be aware of a subject (or act of awareness), or of how he would know that he had found one, and it seems pretty certain that none of them has such an account to offer. Why, then, do they think that they have failed to find a subject by introspection? I think that this is because it seems to them, even before they "look" into themselves (if, indeed, they bother to make a show of looking), that *whatever* they find will inevitably be something other than the subject of awareness. And this is precisely what Figure 2 suggests. This picture is based on the distinction between the perceiver and the perceived, is indeed simply a graphic representation of that distinction, and the perceiver and what it perceives are represented in the picture by distinct things. So as the picture is set up, the perceiver is automatically excluded from what is perceived.

Now in a sense those philosophers who have attacked the notion of a subject and advocated some form of the bundle theory have been attacking the picture of the self given by Figure 2 and Figure 3. Yet they have not *completely* repudiated this picture. It is what they start from. What they have done is to cut off, as it were, the part of the picture that represents the subject of perception or acquaintance, keeping the part that represents the objects of acquaintance. And what remains becomes their picture of a person or self. A person at a particular time becomes, on this picture, simply a collection of objects (i.e., the collection consisting of those feelings, thoughts, images, and so on which the person is said to "have" at that time), and the history of a person becomes just a series of such collections.

Paradoxically, the original picture, even though *it* represents the person as a subject which perceives these objects, seems in the end to suggest that the person is the aggregate of the objects themselves. For the conception that originally suggests

this picture is based on the idea that a person knows the truth of his first-person psychological statements and must know this *from* what he perceives. But these statements, having as their subject the word "I," are regarded as statements about a *person*, namely the person who asserts them. And if they are based *solely* on what is immediately perceived and are known with perfect certainty, it would seem that they must be statements *about* what is immediately perceived. And from this it would follow that they are statements about the objects of perception, not about any subject. If so, then if they are statements about a person, a person cannot be a subject. The view that seems to follow from this picture is the view that to make a statement about a person is simply to make a statement about "mental objects" (feelings, thoughts, images, and so on) and their relations to one another. And this makes it plausible to say that in some sense a person *is* that collection of thoughts, images, and so on that he is said to "have."

The view that a person is a subject seems paradoxical when it is represented as in Figure 2, for this picture seems to suggest that statements about a person (here one is thinking only of first-person statements) are not statements about a subject, from which it seems to follow that a person is not a subject. But the bundle theory is equally paradoxical, and in a similar way. For the picture it gives of a person, the picture that results when we eliminate the subject in Figure 2 and Figure 3, is derived from Figure 1 and Figure 2 and seems to presuppose the very thing that is eliminated from them. The virtue of the bundle theory is supposed to be that it sticks to what is empirically given and does not posit unobservable entities. By constructing the self out of images, thoughts, "perceptions," and so on, the bundle theorist claims to make intelligible the fact that a person *can* have self-knowledge. For the things out of which he constructs the self are, it seems, just the sort of things that can be observed. As Ayer put it,

"The considerations which make it necessary, as Berkeley saw, to give a phenomenalist account of material things, make it necessary also, as Berkeley did not see, to give a phenomenalist account of the self." [36] The paradox appears when one reminds oneself that in order for there to be something that is observed there must be something that observes. If pains, images, and so on are *not* observed then the bundle theory is no more "empirical" than the view that a person is a "substantival ego," and is completely without a point. If they *are* observed then surely there must be something that observes them, and what can that be if not the subject which the bundle theory wants to reject?

Prima facie, then, both the view that there is a subject and the denial of this view are paradoxical. But ways of avoiding both paradoxes have been proposed. Those who say that a self is a subject have sometimes sought to avoid the paradox by maintaining that, contrary to what is represented in Figure 2, the subject of acquaintance is itself an object of acquaintance. What they have done, in effect, is to transform Figure 2 into Figure 6, and Figure 5 into Figure 7. It seems most unlikely that Figure 6 would be the first picture someone would produce if asked to give a schematic representation of what is hap-

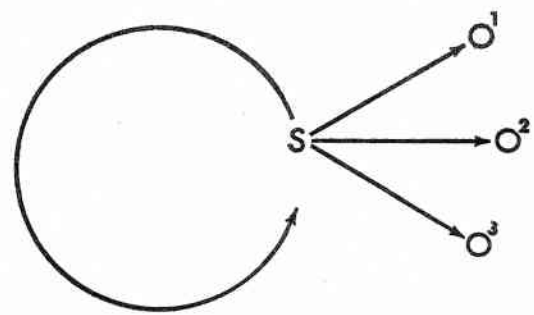

Figure 6

[36] *Language Truth and Logic*, p. 126.

pening in the mind of a person who (for example) has a pain, sees an image, and feels warmth, or that Figure 7 would be the first picture someone would produce if asked to represent what such a person would be aware of. For one thing, the analogy

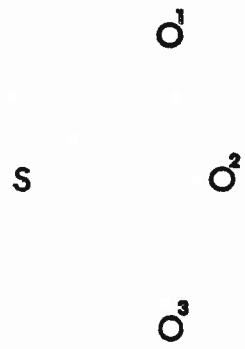

Figure 7

between the awareness of mental objects and visual perception, which makes Figure 2 seem appropriate, breaks down if we represent the subject of awareness as being itself an object of awareness. One cannot see one's eyes (except in a mirror) if one is seeing with them; the place *from* which one sees is necessarily excluded from one's field of vision. The perception of the subject by itself, as represented in Figure 6, is introduced only as a later sophistication, to take care of the difficulties raised above.

The bundle theory, in its more sophisticated forms, adopts a rather different way of avoiding the paradoxical consequences apparently implicit in it. This theory holds that all first-person statements, such as "I have a pain," are statements about "objects" and are not about any "subject." The objection was raised that in order to be an object in the required sense something must be an object *of* perception (awareness, acquaintance),

and that if something is an object of perception, i.e., is perceived, there must be something that perceives it. Instead of denying this (as Hume was perhaps inclined to do) the sophisticated bundle theorist resorts to the notion of analysis. Even the statement that something is perceived, he says, is a statement solely about objects. Although the ordinary forms of expression suggest that perception, or acquaintance, is a two-term relationship, any sentence in which the word "perceive," or a synonym of it, occurs must be analyzable into a sentence in which no such expression occurs and one that is different in form from the original sentence. The assertion that a person perceives a certain object, on this view, will turn out, when analyzed, to be either a statement solely about the "perceived" object itself, to the effect that it exists or has a certain property, or a statement about it and other objects, to the effect that it is related to those objects in a certain way.

On this view, person-referring expressions will disappear in the analysis of psychological statements. And it is a consequence of this view, I think, that it is necessarily false or senseless, rather than contingently false, to say that we are acquainted with the referent of the word "I," or with a subject of experience. For to say this, if the logical construction theory is true, is to misunderstand the use of the word "I"; it is, in Russellian terms, to suppose that "I" is a "logically proper name," when in fact it is an "incomplete symbol." If this is so, however, anyone who "looks into himself" in an attempt to find a referent for the word "I" is rather in the position of someone who looks in his bureau drawer in search of Platonic Forms; he cannot properly be said to be looking for something at all. It follows that the bundle theorist cannot intelligibly claim that his denial that we are acquainted with a subject is based on introspection. This denial, if it is to be used to support the bundle theory, must be based on a priori rather than

empirical grounds. I think that there is a conclusive a priori argument against the possibility of self-acquaintance (if this is regarded as a perception of the "I" that entitles one to make first-person statements). But, as I shall argue in the following chapter, this argument can be directed with equal force against the bundle theory and its account of the nature of self-knowledge.

Three

The Self and the Contents of Consciousness

1. "Why do we regard our present and past experiences as all parts of *one* experience, namely the experience we call 'ours'?" [1] This question of Russell's is both a question about the nature of self-knowledge and a question about the nature of selves. There is the question of how I *know* of my present and past experiences that they are mine (and therefore the experiences of a single person), and there is the question of what *makes* a set of experiences mine (or the experiences of a single person). The second of these questions might also be expressed by asking what it *means* to say that certain experiences are mine (or are the experiences of a single person). Dividing Russell's question in another way, there is the question of how I know, and of what it means to say, that certain *past* experiences are mine (are the experiences of a single person), and

[1] Russell, "On the Nature of Acquaintance," p. 131.