

CRITICAL STUDIES

THINGS AND PERSONS

ALVIN PLANTINGA

RECENT YEARS have seen a resurgence of interest in such traditional problems of metaphysics as the problem of universals, of other minds, of the nature of predication, of the nature of time and space, of whether and in what sense material objects are "basic," and the like. P. F. Strawson's book, *Individuals*,¹ is an important product of that renewed interest. Exciting, able and challenging, the book is also, unhappily, obscure. It is as if Strawson were so eager to pursue the main vision, that he had neither the space nor the inclination to work out carefully the connections between the various things he says or the arguments for his crucial contentions—connections and arguments which are often essential for an understanding of the book as a whole or the main points it makes. Since I am less than confident of my understanding of the book, I offer what follows with diffidence; and my critical remarks are to be taken in the spirit of queries, requests for clarification, rather than as settled criticisms.

Individuals is divided into two parts. In Part I, Strawson tries to show that there is an important sense in which *material objects* and *persons* are *basic particulars*. In Part II "the aim is to establish and explain the connexion between the idea of a particular in general and that of an object of reference or logical subject" (11-12). Part II contains a wealth of interesting material and ingenious argumentation. Most important here, perhaps, is Strawson's attempt to provide criteria for distinguishing subjects from predicates and particulars from universals, and to show a philosophically important link between the subject-predicate distinction and the particulars-universals distinction. What he says

¹ *Individuals. An Essay in Descriptive Metaphysics*, London, Methuen & Co., Ltd. 1959. Unless otherwise noted page references are to that book.

on this head is of considerable interest indeed.² Nevertheless, I shall deal only with Part I. First I shall recount and comment upon the argument for the view that material objects are basic; then I shall state and examine what Strawson has to say about *persons* in more detail.

I

“Material objects are *basic* particulars.” To understand this we must first note what Strawson says about *identification*; for material objects, he holds, are basic with respect to identification. Roughly, to identify a particular about which one makes a remark is to pick it out, conceptually discriminate it from other particulars (speaker identification); and to identify a particular about which someone else has made a remark is to know which particular it is that has been referred to in that remark (hearer identification). In explaining this Strawson speaks of the criteria or *requirements* of identification:

It seems that the general requirements of hearer-identification could be regarded as fulfilled if the hearer knew that the particular being referred to was identical with some particular about which he knew some individuating fact, or facts, other than the fact that it was the particular being referred to. To know an individuating fact about a particular is to know that such-and-such a thing is true of that particular and of no other particular whatever. . . . This, then is the general condition for hearer identification in the non-demonstrative case; and it is obvious that, if a genuine reference is being made, the speaker, too, must satisfy a similar condition. (23)

How is this requirement ordinarily fulfilled? The answer is that the individuating fact known about a particular by someone identifying it is a fact which locates it in “our common spatio-temporal framework” or uniquely relates it to some other particular whose place in the framework is known:

Yet it cannot be denied that each of us is, at any moment, in possession of such a framework—a unified framework of knowledge of particulars, in which we ourselves and, usually, our immediate surroundings have their place, and of which each element is uniquely related to every other and hence to ourselves and our surroundings. It cannot

² Strawson’s discussion is easier to understand if one reads Ramsey’s “Universals” (*Foundations of Mathematics*) in connection with it.

be denied that this framework of knowledge supplies a uniquely efficient means of adding identified particulars to our stock. This framework we use for this purpose: not just occasionally and adventitiously, but always and essentially. (24)

It might be objected that *any* individuating fact about a particular satisfies the requirement laid down for identification; why then the pre-eminence accorded spatio-temporal relations? Strawson replies by disputing the suggestion that one could know an individuating fact about an individual without knowing something about that individual's relations to items whose place in the spatial-temporal framework is known; and he adds that even if the suggestion were made good the achievement would remain "peculiarly useless":

Even if it were possible to satisfy the formal conditions of particular-identification in a way which left the particular completely detached and cut off, as it were, from the general unified framework of knowledge of particulars, the achievement would be a peculiarly useless one. *So long as our knowledge of it retained this completely detached character*, the particular would have no part to play in our general scheme of knowledge; we could for example, learn nothing new about it except by learning new general truths. I do not think we need pursue the question any further; for it is obvious enough that the possibility envisaged, if it is one, plays no significant part in our general scheme of knowledge of particular things. (28; Strawson's italics.)

The argument this far is obviously not tight; but Strawson takes its upshot to be that we can identify particulars only if we employ "a single unified spatio-temporal system." And the next step in the argument is the claim that we can have such a system only if we are able to *re-identify* objects—i.e., able to identify object B encountered at time t_2 as the *same object* as object A encountered at time t_1 . Hence we must have "criteria of re-identification."

Why are criteria of re-identification necessary to our operating the scheme of a single unified spatio-temporal framework for referential identification? . . . Evidently we can sometimes referentially identify a member of the spatio-temporal framework by giving, or being given, its position relative to others. No less evidently we cannot make the identification of every element in the system in this way relative to that of other elements. An immediate answer is that we have no need to, because we can identify some elements by direct location.

But this answer, by itself, is insufficient. For we do not use a different scheme, a different framework, on each occasion. It is the essence of the matter that we use the same framework on different occasions. We must not only identify some elements in a non-relative way, we must identify them as just the elements they are of a single continuously usable system of elements. For the occasions of reference themselves have different places in the single system of reference. We cannot attach one occasion to another unless, from occasion to occasion, we can re-identify elements common to different occasions. (32)

Here we come to the heart of the argument. To say that a given kind of particulars is *basic* is to say something like the following: particulars of kind *a* are basic particulars if and only if we can identify particulars of kind *a* without referring to particulars of any other kind, and there are particulars of some other kind such that we cannot identify particulars of that kind without referring to particulars of kind *a*. (Such things as pains and after images, says Strawson, are examples of particulars that are *not* basic; for in general we can identify a pain only as So and So's pain. Our identification of pains essentially involves a reference to particulars of a different type, namely persons.) And from the premise that we can identify particulars only by relating them uniquely to this "unitary spatio-temporal framework" Strawson infers that material bodies are basic particulars in the above sense:

It seems that we can construct an argument from the premise that identification rests ultimately on location in a unitary spatio-temporal framework of four dimensions, to the conclusion that a certain class of particulars is basic in the sense I have explained. For that framework is not something extraneous to the objects in reality of which we speak. If we ask what constitutes the framework, we must look to those objects themselves, or some among them. But not every category of particular objects which we recognize is competent to constitute such a framework. The only objects which can constitute it are those which can confer upon it its own fundamental characteristics. That is to say, they must be three-dimensional objects with some endurance through time. They must also be accessible to such means of observation as we have; and since those means are strictly limited in power, they must collectively have enough diversity, richness, stability and endurance to make possible and natural just that conception of a single unitary framework which we possess. Of the categories of objects which we recognize, only those satisfy these requirements which are, or possess material bodies—in a broad sense of the expression. Material bodies constitute the framework. Hence, given a certain general feature of the conceptual scheme we possess, and given

the character of the available major categories, things which are or possess material bodies must be the basic particulars. (39)

This argument is extraordinarily puzzling; and indeed Strawson himself seems to distrust it.³ Accordingly he gives a different argument (or perhaps an explication of the same argument): first he points out that if a class of particulars is basic, then particulars of that class must be “publicly observable” (45); he then asks what categories of publicly observable entities are available. As it turns out, there are three such categories: (1) events and processes, (2) states and conditions, and (3) material bodies or things possessing material bodies. Strawson’s argument in essence consists in showing that due to “the contingent limitation of human powers” particulars of groups (1) and (2) do not provide “frameworks of the kind which are at all adequate to our referring needs”; they do not provide frameworks which could be used for identificatory purposes. Material bodies, on the other hand, can and do provide such a framework; hence they and they alone are basic particulars.

Strawson’s argument is unusually difficult to summarize; for one thing its richness and complexity are thereby lost, and with them something of the argument’s strength. And it must be added that it is next to impossible to state the argument (or any important stage of it) precisely; for at every step possible ambiguities and actual obscurities crop up. Nevertheless the above is not an altogether unfair summary: at any rate it represents something of the tenor of the argument. I want to elucidate and comment upon one stage of that argument.

In arguing that our ability to *re-identify* particulars is a necessary condition of our “operating the scheme of a single spatio-temporal framework,” Strawson distinguishes *qualitative* identity from *numerical* or *particular* identity. Though he does not give a *definition* of either, he indicates the distinction he has in mind by an example; and a necessary condition of the numerical identity of individual *a* at time *t* with *b* at some later time *t'*, in Strawson’s

³ “To rest any philosophical position on an argument so general and so vague would be undesirable” (40).

use, is that there be no time between t and t' at which neither a nor b exists (34). To *re-identify* a particular, then, is to identify it as (numerically) the same particular as one encountered on a previous occasion.

Now the distinction between qualitative and numerical identity is easy to apply in the case of objects which are under observation—say two cigarettes of the same brand on the table before me. But very often, of course, a particular is not under observation at every moment of its existence; and very often we judge that particular a at time t is numerically identical with b at t' when there was a time between t and t' at which the particular in question was not under observation. And here we encounter a venerable problem. We *think* we have the right to make the distinction between qualitative and numerical identity in the case of particulars which are not under continuous observation; but *do* we? Suppose I leave my room and return five minutes later: what is my reason for assuming that the desk in the room when I return is numerically identical with the desk that was there when I left? What gives me the right to believe that my desk has enjoyed uninterrupted continuity of existence in my absence? Isn't it true, the skeptic might urge, that in the cases where we are inclined to judge that a at t is numerically identical with b at t' (where there was a time between t and t' at which neither a nor b was under observation) the most we have a right to be sure of is that a and b are *qualitatively* identical?⁴

The first answer might be that we have *criteria* for determining whether in such a case a and b are numerically identical or only qualitatively identical. But, the skeptic replies, why do we take it that the satisfaction of these criteria guarantees numerical identity? For surely it is logically possible in any given case that though the criteria have been fulfilled, a and b are not numerically identical. Surely we can't construct a deductively valid argument for the conclusion that a and b are numerically identical from premisses stating that the criteria for particular identity have been fulfilled. Nor could we construct a good inductive argument for

⁴ Compare Hume's *Treatise*, Part IV, Sections ii, vi.

the same conclusion, for there is no way, independent of the criteria in question, of knowing of temporal continuity of existence in cases of non-continuous observation. Why then do we accept the satisfaction of these criteria as a guarantee of particular identity? All we can be sure of in these cases, declares the skeptic, is qualitative identity.

Here Strawson makes an extremely interesting move; he gives a refutation of the particular-identity skeptic by contending that the skeptic cannot so much as *state his doubts* without tacitly accepting the very distinction he impugns. I quote in full the passage in which this argument occurs:

There is no doubt that we have the idea of a single spatio-temporal system of material things; the idea of every material thing at any time being spatially related, in various ways at various times, to every other at every time. There is no doubt at all that this *is* our conceptual scheme. Now I say that a *condition* of our having this conceptual scheme is the unquestioning acceptance of particular-identity in at least some cases of non-continuous observation. Let us suppose for a moment that we were never willing to ascribe particular-identity in such cases. Then we should, as it were, have the idea of a new, a different, spatial system for each new continuous stretch of observation. . . . Each new system would be wholly independent of every other. There would be no question of *doubt* about the identity of an item in one system with an item on another. For such a doubt makes sense only if the two systems are not independent, if they are parts, in some way connected, of a single system which includes them both. But the condition of having such a system is precisely the condition that there should be satisfiable and commonly satisfied criteria for the identity of at least some items in one sub-system with some items in the other. This gives us a more profound characterization of the skeptic's position. He pretends to accept a conceptual scheme, but at the same time quietly rejects one of the conditions of its employment. Thus his doubts are unreal, not simply because they are logically irresolvable doubts, but because they amount to the rejection of the whole conceptual scheme within which alone such doubts make sense (35. Strawson's italics).

This is a fascinating argument, replete with Kantian overtones. I wish to examine it more closely. But here a note of caution must be sounded; the argument is not easy to state, and I am not altogether confident that what follows is the only plausible interpretation of the passage in which it is given. But at any rate it is *a* plausible interpretation, and an interesting one.

We must be careful to note exactly what the skeptic, as Strawson depicts him, is contending. He is holding, first, that a necessary condition of particular-identity is temporal continuity of existence; and in this Strawson concurs. He is agreeing, secondly, that we do have criteria for particular-identity in cases where temporal continuity of existence is not observed, and that these criteria are in fact often fulfilled. But his skepticism consists in his *questioning* these criteria; in his *doubting* that their fulfillment ever gives us adequate ground for asserting particular-identity. And Strawson's aim is to show that such doubt is radically absurd by showing that a necessary condition of the very meaningfulness of the skeptic's doubt is his unquestioning acceptance of particular-identity in some cases where temporal continuity is not observed. In the following outline of Strawson's argument I shall stick as close as possible to his own words:

- (1) Spatial systems S and S' are sub-systems of the same spatial system if and only if there are some particulars that are members of both S and S'.⁵
- (2) Spatial systems S and S' are different spatial systems if and only if they are not sub-systems of the same spatial system.
- (3) So if we never unquestioningly accepted particular-identity where temporal continuity of existence is not observed, we should have the idea of a different spatial system for each new continuous stretch of observation.
- (4) If spatial systems S and S' are different spatial systems, any doubt as to whether an item in system S is numerically identical with an item in S' is senseless.
- (5) So if we had a different spatial system for each new continuous stretch of observation, any doubt as to whether an item observed in one continuous stretch of observation is numerically identical

⁵ (1) obviously needs amendment. As it stands, it entails that if S and S' lack a common member, then they aren't sub-systems of the same system *even if* there is a third system S'' which shares a member both with S and S'; and Strawson evidently does not wish to accept that consequence. The difficulty can be overcome by (a) defining "neighboring sub-systems" in the way (1) defines "sub-systems of the same spatial system" and (b) defining "sub-systems of the same spatial system" as follows: S and S' are sub-systems of the same spatial system if and only if S and S' are members of a set of neighboring sub-systems such that each member of the set is a neighbor of at least one other member of the set. I leave the reader to make the necessary qualifications in the ensuing discussion.

with an item observed in some other continuous stretch of observation would be senseless.

- (6) Therefore if we never unquestioningly accepted particular-identity where temporal continuity of existence was not observed, we should never be able to raise a sensible doubt as to whether particular *a*, observed at *t*, is numerically identical with particular *b*, observed at *t'*, if there was a time between *t* and *t'* at which neither *a* nor *b* was observed.

And presumably the skeptic is in just the unhappy position proscribed by (6)—i.e., he recommends that we unquestioningly accept particular-identity only where there is observed continuity of existence, yet he still wonders whether in fact there is or isn't particular-identity in cases where continuity of existence is *not* observed.

Is Strawson's argument acceptable? It has more than one difficulty, but the crucial problem here is to evaluate step (3). And here the phrase "continuous stretch of observation" is crucially ambiguous. Strawson may mean by it *either* (1) a period of time during which a single particular is always under observation by *someone or other*, or (2) a period during which a single person is always observing some object or other. And depending upon which of these is intended here, there are two quite different skeptics under attack. One is the moderate skeptic who is doubtful of the particular identity of *a* at *t* with *b* at *t'* only if there was a time between *t* and *t'* at which *no one* was observing either *a* or *b*. The other, the radical skeptic, is doubtful of the particular identity of *a* at *t* with *b* at *t'*, in any case in which *he has not himself* continuously observed the particular which both *a* and *b* name from *t* to *t'*. I am *inclined* to think that it is the moderate skeptic Strawson means to be attacking here. But the text doesn't indicate which he has in mind; hence we must examine the argument under both interpretations of "continuous stretch of observation."

Suppose we take the first interpretation of that phrase. Step (3) may then be rewritten as follows:

- (3') If we never unquestioningly accepted particular-identity where temporal continuity of existence was not observed, then whenever we observed a particular *a* at time *t* and a particular *b* at time *t'* and there was a time between *t* and *t'* at which neither *a*

nor b was observed, we should think that a and b belonged to different spatial systems.

Strawson evidently thinks that steps (1) and (2) of the argument entail (3). Now (2) is the immediate consequence of a *definition* (of "different spatial systems") and is therefore a *necessary* proposition; likewise (1), as a statement of the necessary and sufficient conditions of "system-inclusion" for spatial systems, is presumably necessary. (3), on the other hand, appears to be a *contingent* statement as to what human beings would think or believe in certain circumstances. It is therefore extremely difficult to see how (1) and (2) could entail (3'). What's meant here, I gather, is that (1) and (2) entail that if there were no particular-identity where there was no observed temporal continuity of existence, then there would in fact be a different spatial system for each new stretch of continuous observation; and human beings generally believe (or ought to believe, perhaps) that that proposition is true. About this claim as to what people do or do not think I have nothing to say; the interesting question here is whether what they are said to believe is in fact so, that is, whether in fact (1) and (2) entail

(3'') If there is no particular-identity where there is no observed continuity of existence, then if a particular a is observed at t and a particular b is observed at t' and there is a time between t and t' at which neither a nor b is observed, then a and b belong to different spatial systems.

I said this is the interesting question, because (2), as the consequence of a definition, is unexceptionable and (1) seems at least plausible⁶, while (3'') is utterly incredible. And in fact the entailment doesn't hold, as the following example indicates. Suppose that spatial system S consists in particulars a and c at time t , while system S' consists in particulars b and c at time t' . Suppose further, that though a is observed at t and b is observed at t' , there is a time between t and t' at which neither a nor b is observed. Suppose finally that c is observed from t through t' . Now let us grant (1), (2) and the antecedent of (3''), i.e., that there is no partic-

⁶ Although someone might object that it entails the "relative theory of space," a theory which isn't subscribed to by all philosophers.

ular-identity where there is no observed continuity of existence. The consequent of (3'') is itself a conditional; and the antecedent of that conditional is true of our example. Does it follow that S and S' are *different* spatial systems, i.e., not sub-systems of the same system? Surely not, for *c* is a member of both S and S'; hence by (1) S and S' *are* sub-systems of the same system.

What the example shows is that even if the moderate skeptic accepts (1) and (2) he can be doubtful of or deny particular-identity where temporal continuity of existence is not observed while maintaining quite properly that no violence is thereby done to our "spatio-temporal conceptual scheme." For the condition of inclusion in a single spatial system laid down in (2) is satisfied for spatial systems S_1 to S_n if there is a set of overlapping cases of particular-identity such that for each system S and its successor S' there is at least one particular which is a member both of S and of S'. Hence Strawson's condition for system inclusion is very often fulfilled to the skeptic's satisfaction; it is fulfilled, for example, if there is one particular which is under continuous observation —e.g., the door to Fort Knox. But it can be fulfilled much less exactly; what's necessary is only that for every two successive instants there be a particular under observation at both instants. And the likelihood is that *that* condition has been satisfied as long as there have been users of language. Hence, the moderate skeptic argues, there is no reason at all to think that his doubts impugn the integrity of our "spatio-temporal conceptual scheme." And if so, his position hasn't been shown to be absurd after all.

One more remark about this argument against the moderate skeptic. Suppose he convinced us that we have no justification for ascribing numerical identity where temporal continuity of existence is not observed, and suppose we then refused to predicate numerical identity in any such cases. Suppose further that we came to believe that every so often there were stretches of time during which nothing at all was under observation. According to Strawson we should then no longer believe or be justified in believing that every material body at any time is spatially related to every other material body at every other time. But why should we think so? It seems more than likely that under these circum-

stances we should go on just as before. We should continue to use the same calendars, history books and the like; we should still think that Jerusalem in King Solomon's time was roughly one-hundred-fifty miles from the place where the Red Sea is now; that once there was a forested wilderness where New York City now stands. And in so thinking, we should betray no fundamental misconceptions or confusions. The moderate skeptic's doubts may indeed be empty; but the argument does not seem to show that they are misguided in the way Strawson suggests.

So much for the argument against the moderate skeptic. The situation is a bit different, however, if it is the *radical* skeptic Strawson is after. *He* holds that a person *N* has a right to be sure of the numerical identity of *a* at *t* with *b* at *t'* only if there was no time between *t* and *t'* at which *N* was observing neither *a* nor *b*. And Strawson's attack presumably goes as follows: "If you accept the numerical identity of *a* at *t* with *b* at *t'* only if there was no time between *t* and *t'* at which you were observing neither *a* nor *b*, then you will have a different spatial system for each new period during which you are continuously observing particulars. But if you have a different spatial system for each new continuous stretch of observation, you can't sensibly be in doubt about whether a particular *a* observed in one continuous stretch of observation, is numerically identical with *b* observed in a different stretch; for such a doubt makes sense only if *a* and *b* are *not* in different spatial systems. Hence you cannot sensibly doubt numerical identity in *every* case where you fail to observe continuity of existence; and if you share with the rest of us the conceptual scheme which we do in fact all have (according to which every particular at any time is spatially related to every other at every other time), you must accept particular-identity in at least some cases where you don't observe continuity of existence."

This argument seems considerably stronger than the argument against the moderate skeptic; nevertheless it will not suffice to put the radical skeptic out of business. His rejoinder, I think, will consist in pointing out two things. First, the argument doesn't show that there is any specific case or type of cases in which the skeptic must accept numerical identity in the absence of observed

continuity of existence. At best it shows that if he wonders whether a particular in one of his continuous stretches of observation is numerically identical with a particular in another such stretch, he must believe that there is some particular (or set of temporally overlapping particulars) which persists from the time of the earlier continuous stretch of observation to the time of the later one. But his accepting that belief is not inconsistent with his questioning, in each specific case, the criteria for numerical identity where continuity of existence is not observed. And secondly, the skeptic will point out, what Strawson's argument really shows is that in cases where *a* occurs in one continuous stretch of observation and *b* in another, the doubt is not merely as to whether *a* and *b* are numerically identical; it is also as to whether *a* and *b* are *in the same spatial system*. Nor is he likely to be impressed by Strawson's insistence that we do in fact have a conceptual scheme whereby every particular is spatially related to every other particular. For accepting that scheme consists partly, if Strawson is right, in holding the belief that for any moment *t* and its successor *t'* there is some particular existing both at *t* and *t'*; and the skeptic's claim is just that none of us has good reason to believe that. The fact that most of us *do* believe it is not likely to shake him; he's known all along that most of us (mistakenly in his view) are non-skeptics.

The upshot of the above discussion, I think, is that Strawson fails to refute either the moderate or radical skeptic; and insofar as that refutation is an essential stage in the argument for the basic character of material objects, we must conclude that the prospects for that argument are not bright.

II

Many recent philosophers have said that we predicate mental states—pains, attitudes, emotions, etc.—of others on the basis of their *behavior*. Some philosophers have held that propositions about mental states are in principle *analyzable into* propositions about behavior; that in principle for any proposition ascribing a mental state to some person N, there is a set of statements about N's

behavior which *entails* and is *entailed* by that proposition. Often called philosophical behaviorism, this position has suffered telling criticism. Others have offered a more sophisticated (and more obscure) suggestion to the effect that a man's behavior is our *criterion* for the ascription of mental states to him. Central to that view is an attack upon the supposition, often attributed to Descartes, that the relation between mental state propositions and behavior proposition is in every case merely inductive or contingent; it is argued that it is in general impossible to construct a cogent inductive argument whose premises are propositions about a man's behavior and whose conclusion is a proposition ascribing a mental state to that man. But what distinguishes this view from philosophical behaviorism is the denial, on the part of those who hold it, that mental state propositions are in general analyzable into or deductively related to behavior propositions. And what is puzzling about the view is the contention that though mental state propositions are not, in general, either inductively or deductively related to behavior propositions, behavior is nonetheless the *basis* for third person ascriptions of mental states, and is sometimes an entirely adequate basis—one such that given the behavior in question, there is *no doubt whatever* that the mental state proposition is true. As Strawson puts it, behavior in some cases at least furnishes us with *logically adequate criteria* for the ascription of mental state propositions (105)—where “logically adequate” is not, presumably, to be interpreted in terms of entailment. And this leaves the notion of “logically adequate criterion” shrouded in obscurity.

In the chapter entitled “Persons,” Strawson reiterates this view. He goes beyond it, however, in maintaining, via an argument strongly reminiscent of some of Kant's transcendental arguments, not merely that this “criteriological” relation in fact holds between mental states and behavior, but that its holding is a logically necessary condition of the ascription of mental states either to oneself or to another. As I understand Strawson this is one important aspect of the central thesis of the chapter. That thesis is the contention that the concept of person is *logically primitive*. Precisely what is meant by that? It is not easy to be sure; but

there is reason to think that Strawson's view consists in the conjunction of the following three theses:

- (a) the concept of a person is not to be analyzed, in the Cartesian fashion, as the concept of a "secondary kind of entity in relation to two primary kinds, viz. a particular consciousness and a particular human body" (105).
- (b) the existence of logically adequate behavioral criteria for the ascription of mental states to others is a necessary condition of the ascription of mental states to anything at all (106), and
- (c) there are many cases in which one has an entirely adequate basis for ascribing a mental state to oneself, although this basis is quite distinct from the basis upon which one ascribes that predicate to someone else (107).

These three theses are evidently the crucial contentions of the chapter and the heart of what Strawson has to say about persons. I propose to elucidate (a) and (b), and examine the arguments by which Strawson seeks to establish them, ignoring (c) altogether.

A

The Cartesian analysis of a person, as depicted by Strawson, represents a person as a union of two substances. These two substances are of different logical types: no property significantly predicable of one is significantly predicable of the other. What Strawson is rejecting, then, is the view that a person is really a composite of two substances such that to one of them mental predicates alone are ascribable while to the other only physical or bodily properties can be ascribed. Why does Strawson reject this view? His reasoning begins with the claim that "it is a necessary condition of ascribing states of consciousness, experiences, to oneself, in the way one does, that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself" (99). Secondly, a necessary condition of the ability to ascribe states of consciousness to others is the ability to identify, referentially think about, those things other than oneself to which states of consciousness are to be ascribed. And thirdly, "if the things one ascribes states of consciousness to, in ascribing them to others, are thought of as a set of Cartesian egos to which only private

experiences can, in correct logical grammar, be ascribed . . . , then there is no question of telling that a private experience is another's" (100). The point here is that unless I have a way of distinguishing one particular of a given type from other particulars of that type, I am unable to ascribe any attribute to any member of that type. But if the particulars in question are "inaccessible" to me (in just the fashion that, on the Cartesian view, other selves are inaccessible) and if the attributes to be ascribed are likewise inaccessible in that same way, I don't have any way of distinguishing, in thought, one of the particulars of that type from others of the same type. And if I am unable to identify other subjects of experience, I am unable to predicate states of consciousness of others; hence I am unable to predicate them of myself. The Cartesian position, therefore, entails that no one is able to predicate states of consciousness of anything whatever; but since we clearly *are* able to predicate them of ourselves and of others, Cartesianism is false.

The Cartesian is likely to reply that the argument is an *ignoratio*. For he does not, of course, assert that only "private experiences can, in correct logical grammar" be predicated of selves or mental substances. After all Descartes held that there is a large class of selves such that to each of the members of this class a causal relationship to a physical body can correctly be ascribed. And, the Cartesian might continue, one could identify another subject of experience as "the subject of experience which stands to *that* body in the same way that the subject of my experiences stand to *this* body." But of course bodies other than my own are not inaccessible to me in the way in which another's states of consciousness are. Strawson takes account of this objection to his argument in the following words:

But this suggestion is useless. It requires me to have noted that *my* experiences stand in a special relation to body M, when it is just the right to speak of *my* experience that is in question. That is to say, it requires me to have noted that *my* experiences stand in a special relation to body M; but it requires me to have noted this as a condition of being able to identify other subjects of experiences, i.e., as a condition of my having the idea of myself as a subject of experience, i.e., as a condition of thinking of any experience as *mine*. (101)

Now exactly *why* is the “right to speak of my experience” in question for the Cartesian? To find out, we shall have to formulate the argument against the Cartesian more precisely.

The first premise is explained in a long footnote (99); as stated in the text and explained in that footnote it is susceptible of several interpretations. I shall suggest an interpretation which seems to me likely both because Strawson’s words suggest that he would accept it and because the premise so interpreted is plausible in its own right. But I shall not argue that my interpretation is in fact Strawson’s meaning: such argument, I fear, would be necessarily inconclusive. Since the premise is a specification of a *general* principle governing predication, I shall state it without specific reference to mental substances and states of consciousness:

- (1) A necessary condition of a person N’s ascribing a predicate P to an individual *a* is that N have a conception of the appropriate occasions for ascribing P to individuals *other than a*.
- (2) N’s being able to *identify* individuals other than *a* is a necessary condition of N’s having a conception of the appropriate occasions for ascribing P to individuals other than *a*.

The crucial word here is “identify.” We should recall that the general requirement, with respect to identification, is the following:

- (3) N can identify a particular only if N knows some individuating fact about that particular (p. 23, 181-183).

We are now in a position to try to understand why Strawson takes it that the Cartesian cannot identify other subjects of experience by means of some such locution as “the subject of experience that stands in the same relation to that body as the subject of my experiences stands to this body.” One is immediately struck by the thought that perhaps Strawson is here using “necessary condition” in such a fashion that “N’s doing *a* is a necessary condition of N’s doing *b*” entails that N must do *a before* he can do *b*. At any rate on this interpretation Strawson’s argument against the Cartesian would be valid. For premises (1) and (2) interpreted in this way entail that one must be able to identify other subjects of experience *before* one can predicate experiences of himself, i.e., before one can think of or refer to his own experiences

at all. And one can identify other subjects of experience in the way the Cartesian suggests only if one *is already able* to predicate experiences of himself, i.e., already able to recognize certain experiences as his own. But this is not, presumably, Strawson's meaning. For (1) and (2) state *general conditions* of predication. And under the present interpretation these conditions are logically incapable of fulfillment, not merely with respect to discourse in the Cartesian mode about other minds, but with respect to any predication of any property of any individual whatever. For under the interpretation in question, the above three principles come to the following: before I can predicate any property of any individual I must have a conception of the appropriate occasions for predicating that property of other individuals (1). Before I can have a conception of the appropriate occasions for predicating that property of other individuals I must be able to *identify* other individuals (2). And being able to identify an individual entails knowing some individuating fact about that individual (3). But if I know an individuating fact about an individual I am already able to predicate a property of that individual, the property, namely, which distinguishes it from other individuals. Hence before I am able to predicate any property of any individual I must be able to predicate some property of some individual; and it is accordingly logically impossible for me to predicate any property of any individual.

Let me try to state this argument a bit more precisely. In the long footnote I referred to above, Strawson says that a predicate is correlative with a range of distinguishable individuals of which it can be significantly, though not necessarily truly, affirmed. We may therefore say that every predicate *determines* a range of individuals—those individuals, namely, of which it can be significantly affirmed (and of course different predicates may determine the same range). Let us consider any predicate P and the range R of individuals it determines. We may restate Strawson's three principles as follows:

- (1) Before a person *N* can predicate *P* of any individual *a* of range *R*, *N* must have a conception of the appropriate occasions for predicating *P* of other members of *R*.⁷
- (2) Before *N* has a conception of the appropriate occasions for predicating *P* of other members of range *R*, *N* must be able to identify other members of *R*.
- (3) If at time *t* *N* identifies any particular *b*, *N* knows some individuating fact about *b* at *t*.

But if at time *t* *N* knows some individuating fact about *b*, then at *t* *N* knows that a certain predicate *Q* applies to *b*.⁸ So if at *t* *N* is able to identify other members of *R*, he must, at *t*, have a conception of the appropriate circumstances for predicating *Q* of individuals of range *R*. Now either *P* and *Q* are the same predicate or they are not. If they are the same, the (absurd) consequence is that *N* has a conception of the appropriate circumstances for predicating *P* of individuals of *R* before he has that conception. If *P* and *Q* are different, on the other hand, a vicious infinite regress ensues: for before *N* has a conception of the appropriate occasions for predicating *Q* of members of *R* he must be able to identify members of *R*. But whenever he is able to identify members of *R* he already has a conception of the appropriate occasions for predicating some third predicate *O* of members of *R*. And so on. The upshot seems to be that under this interpretation of Strawson's premises it is logically impossible for anyone to ascribe any attribute to anything.

If the above is correct, we do well to look for a different interpretation of Strawson's objection to the Cartesian. And a natural suggestion is that Strawson means, not that *N*'s ability to identify other subjects of experience must be temporally prior to his ability to predicate experiences of them, but only that at any time at which *N* is able to predicate experiences of others, he must be able to identify them. Premises (1) and (2) then turn out as follows:

⁷ I take Strawson's expression "having a conception of the appropriate occasions for predicating *P*" as roughly equivalent to "knowing the meaning of *P*."

⁸ "To know an individuating fact about a particular is to know that such-and-such a thing is true of that particular. . . ." (23)

- (1b) If N can predicate states of consciousness of himself at time t N has, at t , a conception of the appropriate occasions for predicating states of consciousness of others.
- (2b) If N has a conception of the appropriate occasions for predicating states of consciousness of others, at t , N is able, at t , to identify other subjects of experience.

The discussion so far has centered about Strawson's claim that one who held a Cartesian position could not make a certain suggestion about the identification of other subjects of experience—the suggestion that another subject of experience could be identified as, e.g., “the subject of experience that stands in the same causal relation to *that* body as the subject of my experiences stands to *this* one.” But now the question becomes acute: why *can't* the Cartesian stick with this suggestion? I should think he could comfortably agree to (1b) and (2b) while maintaining the thesis about identification Strawson denies him. With respect to (1b), for example, he might say the following: if N can predicate mental states of himself at time t , then at t N knows (or believes, thinks, etc.) that for any mental state M he predicates of himself, if he were to observe a correlation between a state B of his body and the presence of M, then any occasions on which he observes a body similar to his in state B is an appropriate occasion for predicating M of another subject of experience. And with respect to (2b) he might claim to be able to identify the subject of experience in question in just the way Strawson prohibits. I quoted above (p. 508) the argument Strawson produces at this point; but under the present interpretation of “necessary condition” that argument turns out to be nothing more than a restatement of principles (1b) and (2b).

We might be tempted to object that if the Cartesian were right *no one could ever know* that there were subjects of experience connected with bodies other than his own. But that would be to change the subject altogether; Strawson has maintained not merely that on Cartesian principles no one could know for sure that there were other subjects of experience (with that the Cartesian might happily agree), but that no one could so much as predicate experiences, states of consciousness, either of others or of *himself*. And if I have interpreted his argument aright, it

doesn't substantiate this claim. We seem to be left with the bare assertion that on the Cartesian view it is impossible to see how anyone could come by the concept of selfhood at all, or how anyone could form the idea of a subject of experience.

B

The second part of the thesis that the concept of a person is primitive has to do with what I called the "criteriological" relationship between behavior and mental states. In rejecting behaviorism, Strawson is rejecting the contention that every third person proposition ascribing a mental state is logically equivalent to a set of propositions which describe behavior. On the other hand, Strawson is emphatic in denying that the relation between mental-state ascribing propositions and behavior propositions is merely inductive or contingent. What, exactly, this relation is, is left obscure in Strawson's explanation as in every other I know of. But in any event, it seems to be held that if A is the criterion for B, then given A, there is no legitimate doubt that B is the case. And what Strawson seeks to establish is not merely that this criteriological relationship does in fact connect behavior and mental states, but that its connecting them is a necessary condition of our being able to predicate mental states either of ourselves or of others; he offers a "transcendental deduction" of the criteriological relation. I quote the argument in full:

Clearly there is no sense in talking of identifiable individuals of a special type, a type namely, such that they possess both M-predicates and P-predicates,⁹ unless there is in principle some way of telling, with regard to any individual of that type, and any P predicate, whether that individual possesses that P-predicate. And in the case of at least some P-predicates, the ways of telling must constitute in some sense logically adequate kinds of criteria for the ascription of P-predicates. For suppose in no case did these ways of telling constitute logically adequate kinds of criteria. Then we should have to think of the relation between the ways of telling and what the P-predicate ascribes, or a part of what it ascribes, always in the following way: we should have to think of the ways of telling as *signs* of the presence, in the

⁹ A P-predicate is one which ascribes a state of consciousness, or implies the possession of consciousness of the part of that to which it is ascribed.

individual concerned, of this different thing, viz., the state of consciousness. But then we could only know that the way of telling was a sign of the presence of the different thing ascribed by the P-predicate, by the observation of correlations between the two. But this observation we could each make only in one case, viz., our own. And now we are back in the position of the defender of Cartesianism who thought our way with it was too short. For what, now, does 'our own case' mean? There is no sense in the idea of ascribing states of consciousness to oneself, or at all, unless the ascriber already knows how to ascribe at least some states of consciousness to others. So he cannot argue in general 'from his own case' to conclusions about how to do this; for unless he already knows how to do this, he has no conception of his own case, or any case, i.e., any subject of experiences (105-106).

Let me try to state this argument a bit more precisely. In order to simplify matters I shall specify the argument with respect to a certain kind of mental state, pain, and the "ways of telling" appropriate to that state which I shall call P-behavior. Strawson's argument may then be outlined as follows:

- (1) If P-behavior is not a logically adequate criterion for the ascription of pain, P-behavior is a *sign* of pain.
- (2) If a person N knows at time t that P-behavior is a sign of pain, then at some prior time $t - 1$ N observed a correlation between P-behavior and pain in his own case.
- (3) If at $t - 1$ N noted a correlation between pain and P-behavior in his own case, then N was able to predicate pain of himself at $t - 1$.
- (4) If at $t - 1$ N was able to predicate pain of himself, then at $t - 1$ N knew how to predicate pain of others.
- (5) If at $t - 1$ N knew how to predicate pain of others, then at $t - 1$ N knew that P-behavior is a sign of pain.

Step (6) would be a repetition of (2) with " $t - 1$ " substituted for " t " and " $t - 2$ " for " $t - 1$ ". Steps (7), (8) and (9) would parallel (3), (4) and (5); step (10) would again parallel (2) with the substitution of " $t - 2$ " for " t " and " $t - 3$ " for " $t - 1$ "; and so on. The conclusion Strawson draws, of course, is that if P-behavior were not a logically adequate criterion for the ascription of pain, no one could ascribe pain to anyone else; and hence no one could ascribe pain to himself either. Once again there are various difficulties in Strawson's argument. I shall concentrate upon one which seems to me central. There is an ambiguity in the phrase "knows how to predicate pain of others" as it occurs

here. Step (4) in this argument is clearly intended to be identical with one of the premises we encountered in Strawson's argument against the Cartesian—the premise, namely, that was step (2) in that argument. (I shall call that premise the Predication Principle.) Accordingly (4) ought to be taken as

- (4a) If at $t - 1$ N was able to predicate pain of himself, then at $t - 1$ N had a conception of the appropriate occasions for predicating pain of others,

and (5) ought to be rewritten as

- (5a) If at $t - 1$ N had a conception of the occasions appropriate for ascribing pain to others, then at $t - 1$ N knew that P-behavior is a sign of pain.

But it is not at all clear that (5a) ought to be accepted. Suppose that in accordance with step (2) N could know that P-behavior is a sign of pain only if he had observed a correlation between P-behavior and pain. At time $t - 2$, *before* he has observed any such correlation, N might make the following declaration: If I observe a connection between my pain and P-behavior on my part at $t - 1$, then any occasion after $t - 1$ on which I observe P-behavior which is not mine will be an appropriate occasion for predicating pain of someone else. If he were to make such a declaration, he would "have a conception of the circumstances appropriate for predicating pain of someone else," although he wouldn't then know that P-behavior is a sign of pain, for he wouldn't have made the requisite observation. Hence it seems that step (5) cannot be accepted.

Now for a couple of objections to my refutation. First, we should note that Strawson distinguishes a weaker and a stronger version of the Predication Principle (99). He claims that his arguments require only the weaker version, which is the one I have been ascribing to him and under which in fact the arguments are not sound. The stronger version of that principle asserts that a necessary condition of one's being able to predicate mental states of himself is that one *in fact* predicate mental states of others (as opposed to merely having a conception of the appropriate occasions for so predicating them). And someone might say that Strawson's argument can be rescued by employing the stronger

version of the Predication Principle. Here we must again remember the two possible interpretations of "necessary condition": the stronger principle *may* be asserting that before N can ascribe pain to himself he must have ascribed it to someone else. So taken, the principle has little to recommend it, for it is supposed to be a specification of a *general* principle about predication; but that general principle would then be the absurd assertion that before I can predicate any property of any one thing I must have predicated that property of something else. The other sense of "necessary condition" is not easy to apply to the Predication Principle in its stronger formulation, but presumably it would come to something like this: I can't have predicated pain of myself before I have predicated it of others—i.e., any time at which I predicate a mental state of myself is a time before which or at which I predicate that state of someone else. And the general principle of which this is a specification would be: Any time at which N predicates a property P of individual *a* is a time at which or before which N predicates that property of some other individual. This in turn entails that the first occasion on which anyone predicates a property of anything is one on which he predicates that property of *at least two things*. I take it that the principle so stated is utterly gratuitous and unworthy of further consideration. But a little reflection will satisfy anyone who is interested that the declaration I ascribed to N at $t-2$ can be reformulated so as to be consistent even with that peculiar version of the Predication Principle.

Second: it might be objected that it is ridiculous to suppose that N could make the declaration I ascribed to him before he had so much as noted a correlation between pain and P-behavior in his own case. And of course, the objector is right; it *is* ridiculous. Doubtless the envisaged state of affairs is psychologically impossible. But the objection is irrelevant; for Strawson's view is that it is a *necessary truth* that if "ways of telling" were never logically adequate criteria for third person mental state propositions, then no one could predicate mental states of anything whatever. Hence it is required only that it be *logically* possible for N to make the declaration I ascribed to him. The thesis that behavior is induc-

tively or contingently related to mental states is not a piece of genetic psychology.

Finally, it might be said that if N could make the declaration I ascribed to him, then he must have learned what pain is "from his own case," and that it is logically impossible for anyone to learn the meaning of "pain" from his own case. Granted that premise, Strawson's argument would be perfectly sound. And no doubt Strawson would accept the premise; indeed, he seems to believe that the argument we have been considering *proves* that contention. But in fact his argument, if I have grasped it, *pre-supposes* rather than *proves* the contention that it is logically impossible to learn the meaning of "pain" from one's own case. And of course that contention is far from obvious; an argument for it is required. An important question, therefore, for the view that the relation between pain and behavior is criteriological, is the question whether a cogent argument can be given for the principle that one can't learn the meaning of words like "pain" from one's own case.

I think it is appropriate to mention that at least two sorts of argument for this principle have been offered. One line of argument is the following: if one could learn what pain was from his own case, one could have a "private language"; but private languages are in principle impossible.¹⁰ I shall not comment upon this argument here; suffice it to say that neither of its contentions is at all easy to establish and both require consideration of difficult and complex issues. The other sort of argument has also been suggested by Malcolm. Commenting upon a passage from Wittgenstein's *Investigations*, he says

The argument that is here adumbrated is, I think, the following: If I were to learn what pain is from perceiving my own pain, then I should, necessarily, have learned that pain is something that exists only when *I* feel pain. For the pain that serves as my paradigm of pain (i.e., my own) has the property of existing only when *I* feel it. That property is essential, not accidental; it is nonsense to suppose that the pain I feel could exist when I did not feel it. So if I obtain my conception of pain from pain that I experience, then it will be part

¹⁰ See N. Malcolm, "Wittgenstein's *Philosophical Investigations*," *Philosophical Review*, October, 1954, pp. 530-534.

of my conception of pain that *I* am the only being that can experience it. For me it will be a *contradiction* to speak of *another's* pain.¹¹

I shall conclude with a comment on Malcolm's argument. Its conclusion—that it is impossible to learn what pain is from one's own case—is just the one needed for Strawson's transcendental deduction of criteria. One might therefore say that Strawson's argument *presupposes* Malcolm's argument, or one like it; his own argument utterly fails to substantiate the claim that the relation between behavior and mental states is criteriological. Malcolm's argument, however, seems to employ something like the following principle:

If all the objects from which I get my concept of C have property P *necessarily*, then for me "All C's have P" will be a necessary proposition.

So stated, the principle is objectionable. For objects don't have necessary properties *merely as objects*, so to speak; they have them only under or relative to certain descriptions. A cow necessarily has the property of being female—i.e., it is a necessary truth that if Bessie is a cow, then Bessie is female. But it's not a necessary truth that Bessie is female; nor that the object named by "Bessie" is a female kine. Consider any proposition of the form: It is a necessary truth that object A has property P. It is clear that "A has P" does not entail "A exists"; for if it did, "A exists" would be a necessary proposition; and propositions asserting that something or other exists are not, in general, necessary. Hence the proposition must be reformulated: It is a necessary truth that anything which is an A has P. But now it becomes clear that "A" is functioning as a *description*. And the question whether in fact "Anything which is an A has P" is necessary will turn on which description is to replace "A" and what property "P" denotes. Before we can understand and apply Malcolm's principle, therefore, we must be told what description it is, of the pains from which I get my concept of pain, under which it is necessarily true that all those pains are felt only by me.

¹¹ *Op. cit.*, p. 538.

Now at this point the following suggestion naturally arises:

If the proposition that object A was one of the objects from which I got my concept of C entails the proposition that A has property P, then my conception of C is such that "All C's have P" is a necessary proposition.

So reformulated, the principle escapes the objection and entails that no one can learn the meaning of "pain" from his own case; for all the pains from which one could so learn what pain is are necessarily one's own. But so reformulated the principle won't do at all: the proposition that A is one of the objects from which I got my concept of C entails that I have *perceived* A; and hence *any* concept I get "from objects" will be such that it necessarily applies only to things I have perceived.¹² Of course, there may be other formulations of Malcolm's principle that are satisfactory. But what is clear is that this argument requires clarification and explanation.

The present status of the view that the relation between mental states and behavior is criteriological, therefore, seems to me to be this. What is needed above all is explanation and development of the two kinds of argument Malcolm discusses in "*Wittgenstein's Philosophical Investigations*." Pending such explanation the view in question remains an interesting but unsubstantiated speculation.

My comments upon Strawson's book have been mainly critical and destructive. I want to point out, in conclusion, that one *compliments* a book by subjecting it to careful, searching criticism. Strawson's book eminently deserves and richly repays that sort of compliment.

Wayne State University.

¹² I owe this point to Robert C. Sleigh, Jr. I take this opportunity to record my indebtedness to him and to my colleagues Edmund Gettier and Hector Castaneda, and to Nicholas Wolterstorff of Calvin College, for valuable criticism and discussion on several points in this paper.