# THE CONCEPT OF A PERSON*

A. J. Ayer
*Oxford University*

## *1*

THE PROBLEMS which I intend to discuss are excessively familiar to students of philosophy. They are concerned with persons in the broad sense in which every individual human being can be counted as a person. It is characteristic of persons in this sense that besides having various physical properties, including that of occupying a continuous series of spatial positions throughout a given period of time, they are also credited with various forms of consciousness. I shall not here try to offer any definition of consciousness. All I can say is that I am speaking of it in the ordinary sense in which, to be thinking about a problem, or remembering some event, or seeing or hearing something, or deciding to do something, or feeling some emotion, such as jealousy or fear, entails being conscious. I am not at this stage committting myself to any view about the way in which this notion of consciousness should be analyzed.

The first question which arises is how these manifestations of consciousness are related to the physical attributes which also belong to persons. The answer which I think would still be most acceptable to common sense, at least when it is made to consider the question in these terms, is that the relation is contingent, not logical, but only factual. In philosophy this

* This essay appeared in A. J. Ayer's *Concept of a Person* (London: Macmillan; New York: St. Martin's Press, 1963).

view is mainly associated with Descartes; if he did not originate it, he put it forward in the clearest and most uncompromising way. The view is that a person is a combination of two separate entities, a body and a mind or soul. Only the mind is conscious; the physical properties which a person has are properties of his body. The two entities are separate in the sense that there is no logical connection between them. It is conceivable that either should exist without the other; that is, there is no contradiction in supposing that a person's mind exists in some other body, or apart from any body at all, and equally none in supposing that a person's body is animated by some other mind, or not by any mind at all. This does not, however, exclude the possibility of there being causal connections between them; so that even if they are separable in principle, there may still be grounds for holding that they are inseparable in fact. Descartes himself prejudged this question by defining the mind as a substance, which implied, in his usage, that its existence was causally as well as logically independent of the existence of the body. But this view that the mind is a substance is not entailed by the view that mind and body are logically distinct. It would be compatible with this sort of dualism to reject the notion of mental substance altogether and conceive of the mind, in Humean fashion, as a series of experiences.

Whatever may be the attractions of this dualistic view for common sense, the tendency of philosophers has been to try to replace it by some form of monism. Thus Berkeley, who held that physical objects were collections of sensible qualities which were dependent for their existence upon being perceived, and Hume, who saw no grounds for holding that anything existed but sensory impressions and the ideas which copied them, may both be regarded, in their different ways, as having tried to effect the reduction of body to mind. In more recent versions of this type of theory, such as those developed by William James and Bertrand Russell, the sense-

data and images, which are taken as fundamental, are held to constitute a kind of neutral stuff, itself neither mental nor physical, out of which both mind and matter are to be constructed. Conversely, it was held by Hobbes, in opposition to Descartes, that there was no need to postulate the existence of minds in addition to bodies; conscious states and activities could be attributed to the body itself. And modern philosophers, like Ryle and Carnap, have argued that it is a mistake to think of conscious states and processes as ghostly inhabitants of a private mental stage; statements about people's mental life are reducible to statements about their physical constitution, or their actual and potential behaviour.

With all of these theories, except perhaps the last, there is a corresponding problem of personal identity. On any dualistic view an account is required of the way in which the mind is lodged in the body. Could there be more than one mind in a single body? Could the same mind dwell in more than one body, at the same or at different times? If the relation is one to one, how are its terms paired off? How is it decided which mind goes with which body? The most plausible answer is that they are causally connected in some special way, but it is not easy to see how this connection is to be defined. If the mind is regarded as a substance, the question arises how such a substance could ever be identified. If it is regarded as a collection of experiences, there is the problem, to which Hume himself confessed that he could see no answer, of showing how the collection is united. What is it that makes a given experience a member of one such collection rather than another? With any view of this type, there is also the problem of identifying the experiences themselves. In the ordinary way, we identify experiences in terms of the persons whose experiences they are: but clearly this will lead to a vicious circle if persons themselves are to be analyzed in terms of their experiences.

An argument in favor of the physicalistic type of monism is

that these difficulties are avoided. At least there is then no special problem of personal identity. The criteria for the identity of persons will be the same as those that determine the identity of their bodies; and these will conform to the general conditions which govern the identity of all physical objects of a solid macroscopic kind. It is primarily a matter of spatio-temporal continuity. Moreover, if persons can be equated with their bodies, there will no longer be any need to specify how minds and bodies are correlated. Once it is shown how states of consciousness can be ascribed to bodies, this problem will have been solved. But whether this can be shown is itself very much an open question. It is obvious that any view of this type encounters very serious difficulties; and it is not at all so clear that they can be satisfactorily met.

What is common to all these theories is the view that the concept of a person is derivative, in the sense that it is capable of being analyzed into simpler elements; they differ only about the character of these elements and the way in which they are combined. But this premise itself has recently been challenged. In his book *Individuals,* Mr. P. F. Strawson has attempted to prove that the concept of a person is a primitive concept; and what he means by this is just that it is not analyzable in any of the ways that we have outlined.[1] Not everything that we want to say about persons can be construed as a statement about the physical objects which are their bodies; still less when we refer to persons are we referring to mental substances, or to collections of experiences. Neither, in Mr. Strawson's view, can it be maintained that persons are compound; that they are the product of two separate entities, or sets of entities, one the subject of physical characteristics and the other the subject of consciousness. He holds, on the contrary, that the subject to which we attribute the properties which imply the presence of consciousness is literally identical with that to which we also attribute

physical properties. And if we ask what this subject is, the only correct answer is just that it is a person.

Mr. Strawson's main reason for rejecting dualism is, in his own words, that "the concept of the pure individual consciousness—the pure ego—is a concept that cannot exist; or, at least, cannot exist as a primary concept in terms of which the concept of a person can be explained or analyzed. It can exist only, if at all, as a secondary, non-primitive concept, which itself is to be explained, analyzed in terms of the concept of a person."[2] It might be thought that this would not affect the dualist who rejects the notion of the pure ego, and thinks of the conscious subject as a collection of experiences, but in fact the reasons which Mr. Strawson has for denying that there can be a primary concept of the pure ego apply equally to any idea of a non-physical subject of consciousness, whatever its composition may be thought to be.

His argument runs as follows: The first premise is "That it is a necessary condition of one's ascribing states of consciousness, experiences, to oneself, in the way that one does, that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself."[3] Now this is understood to imply that one ascribes experiences to others in exactly the same sense as one ascribes them to oneself. It excludes the view, which has been held by some philosophers, that the statements which a person makes about the experiences of others are to be analyzed quite differently from the corresponding statements that he makes about his own, that whereas in his own case he is to be understood as speaking literally, what he says about the experiences of others can only be construed as a reference to their behavior. But when one talks about the experiences of another person one cannot be attributing them to a pure consciousness; neither is it possible to regard the subject of one's statement simply as a collection of experiences. The reason for this is that in the case of another person neither the pure consciousness nor

the collection of experiences would be things that one could have any means of identifying. But if our attributions of experiences to others cannot be understood in this way, neither can our attributions of experiences to ourselves: this follows from the principle that the same analysis must be applied to both.

An important consequence of this argument, if it is sound, is that we must give up the argument from analogy on which many philosophers have relied as a justification for believing in the existence of other minds. It is tempting to think that one can come by the idea of one's own experiences through introspection, observe that in one's own case experiences of certain kinds are characteristically associated with certain forms of behavior, and so, when one observes other people behaving in similar ways, infer that they are having similar experiences. Even if one does not acquire the belief that there are other minds in this fashion, it may still be the ground for holding that the belief is rational. This reasoning has, indeed, met with various objections. Assuming that it is logically impossible for anyone directly to observe what goes on in another person's mind, some philosophers have maintained that this is not a valid argument from analogy; for they hold that no inductive argument can give us any reason to believe in the existence of something which could not even in principle be observed. Others who think that this difficulty can be overcome find fault with the argument because its basis is so weak. As Wittgenstein put it, "How can I generalize from the *one* case so irresponsibly?" The novelty of Mr. Strawson's attack lies in his refusal even to allow the argument to start. If my knowing how to ascribe experiences to others is a necessary condition of my being able to ascribe them to myself, then, Mr. Strawson suggests, the argument begins by presupposing what it is intended to justify.[4]

Moreover, even if, without consideration of others, I could initially distinguish what are in fact my experiences from the

body with which they are associated, and this body from other bodies, this still would not give me any ground, in Mr. Strawson's view, for supposing that any of these other bodies were "owned" by subjects who also had experiences. I should discover empirically that certain feelings occurred when this body was acted on by certain stimuli, and that they did not occur when other bodies were so treated. But all that could ever be in question would be the presence or absence of experiences of my own. And even this goes too far, if it only makes sense for me to talk of my experiences in contradistinction to those of other people. To some extent this argument was anticipated by G. E. Moore, who maintained in his essay on "The Nature and Reality of Objects of Perception" that if one assumes with Berkeley that the objects of perception exist only so long as one is perceiving them, then no reasoning by analogy could give one any ground at all for ascribing experiences to other people; the most that it could possibly authorize would be a belief in the existence of unconscious experiences of one's own. But for Moore this was just an argument against idealism; he thought that if the objects of perception were allowed to be physical bodies which existed independently of our perceiving them, then we could rely upon analogy as a ground for believing that some of these bodies were inhabited by minds like our own.[5] Mr. Strawson, on the other hand, holds that even if we grant this premise about the objects of perception, there is still no basis for the argument from analogy.

It might be thought that this line of reasoning would result in the elimination of anything but the body as a possible subject of consciousness, but Mr. Strawson does not take this view. He does not in fact discuss the thesis of physicalism, according to which statements about experiences are transformable into statements about physical occurrences, but instead goes on to examine a hybrid theory which he calls the "no-ownership" doctrine of the self. This is the theory that

the only sense in which experiences can significantly be said to have an owner is that they are causally dependent upon the state of some particular body. It is perhaps misleading to call this a "no-ownership" theory, since in the sense which it allows to ownership, it does not imply that any experiences are unowned. Mr. Strawson's reason for so calling it is not so much that this sense of ownership is Pickwickian as that it does not yield a guarantee that experiences are private property. For he holds that it must on this theory be regarded as a contingent fact that the experiences which a person owns are causally linked to his body and not to some other body instead.

Mr. Strawson's objection to this theory is that it is incoherent. The proposition which it tries to state is that, with respect to any given person, all his experiences are dependent upon the state of his body; and this proposition is supposed to be contingent. But how are his experiences to be identified? In accordance with what principle are my experiences classified as mine? If the answer is that they are just those experiences which are causally dependent, in the requisite way, upon the state of this body, then the proposition that all my experiences are causally dependent upon the state of this body becomes analytic; it is just a way of saying that all the experiences which are causally dependent upon the state of this body are causally dependent upon the state of this body. But what the theory requires is that this proposition be contingent. And since it admits no other way of identifying a person's experiences, the consequence is that it defeats itself.

It is also true of physicalist theories that they admit no other way of identifying a person's experiences than by identifying his body. It might, therefore, be thought that they too were exposed to Mr. Strawson's argument, especially as it seems to be a contingent proposition that some particular body is the body of such and such a person. But here the physicalist has an effective answer. He can argue that the

reason why this proposition is contingent is just that it presupposes that the body in question has been independently identified, either ostensively or by some other form of description. If we have identified a person, it follows on his view, that we have identified that person's body; but the converse need not hold. Neither is there any further problem for the physicalist about the identification of experiences, since he maintains that statements about a person's experiences are logically equivalent to statements about the condition or movements of his body. As we shall see, there are serious objections to any view of this kind; but on this score at least, it is not incoherent. We may, therefore, conclude that Mr. Strawson's argument is not fatal to theories of this type. Whether it is fatal even to the "no-ownership" theory is one of the questions that we shall have to consider later on.

Physicalist theories are based on a consideration of the way in which we ascribe experiences to other people. Since our only ground for this proceeding is our observation of their physical condition and behavior, it is assumed that this is all that we can be referring to. Then, on the assumption that we must mean the same by the ascription of experiences to ourselves as we do by ascribing them to others, it is inferred that even when we speak about our own experiences we are referring to our physical condition or behavior. To dualists, on the other hand, it is evident that when we speak about our own experiences we are not referring to the physical manifestations by which other people may be made aware of them. Our knowledge of our own experiences is of an entirely different character. So again assuming that when we speak about the experiences of others we must mean the same as we do when speaking about our own, they infer that the physical events on which we base our attributions of experiences to others are signs of these experiences and not to be identified with them. They hold that we have direct knowledge of our own experiences; but that such knowledge as we can have of

the experiences of others is only inferred from their physical manifestations. And then they are faced with the problem how these inferences can be justified.

In maintaining that the concept of a person is logically primitive, Mr. Strawson hopes to secure the advantages and at the same time avoid the difficulties of both these lines of approach. He admits that the basis on which we ascribe experiences to others is different from that on which we ascribe them to ourselves, but he denies that, in the case of other people, we are reduced to making an inductive inference. There is not merely a factual connection between certain physical events and the experiences which they are understood to manifest. It is true that when we ascribe experiences to others we do not simply mean that they are in such and such a physical condition or that they are behaving, or disposed to behave, in such and such ways. These are just the criteria by which we determine that they are having the experience in question. But the point is that these criteria are "logically adequate." In our own case, we do not rely on these criteria. Our knowledge of our own experiences is not obtained by observation. But this does not mean that the sense in which we ascribe experiences to ourselves is in any way different from that in which we ascribe them to others. On the contrary, it is a necessary feature of predicates which imply that the subject to which they are attributed is conscious, "it is essential to the character of these predicates, that they have both first and third person ascriptive uses, that they are both self-ascribable otherwise than on the basis of observation of the behavior of the subject of them, and other ascribable on the basis of behavior-criteria."[6] If we did not understand the use of predicates of this kind, we should not possess the concept of a person. For persons are essentially the subjects to which such predicates are attributed.

But how, to echo Kant, are such predicates possible? Or, as Mr. Strawson puts it, "What is it in the natural facts that

makes it intelligible that we should have this concept (of a person)?"[7] He does not attempt to answer this question in any detail, but he does suggest that if we are looking for an answer we should begin by directing our attention to predicates which are concerned with human action. He thinks that it is easier to understand how we can see each other, and ourselves, as persons, "if we think first of the fact that we act, and act on each other, and act in accordance with a common human nature." His reason for thinking this is that a study of the ways in which we do things should rid us of the belief "that the only things we can know about without observation and inference, or both, are private experiences." In cases of intentional action, we also have knowledge, not based on observation and inference, about the present and future movements of our bodies. Not only that but predicates which refer to forms of action do so as a rule "while not indicating at all precisely any very definite sensation or experience." The result of this is that although we ascribe such predicates to others on the basis of observation and do not in general ascribe them on this basis to ourselves, we find it much easier in their case than in the cases where there is a reference to some distinctive experience to recognize that what is attributed on these different bases is nevertheless the same.

The suggestion that persons are to be distinguished in the first instance by their capacity for action has also been put forward by Professor Hampshire in his recent book on *Thought and Action.*

> The deepest mistake in empiricist theories of perception, descending from Berkeley and Hume, has been the representation of human beings as passive observers receiving impressions from "outside" of the mind where the "outside" includes their own bodies. In fact, I find myself from the beginning able to act upon objects around me . . . I not only perceive my body, I also control it; I not only perceive external objects, I also

> manipulate them. To doubt the existence of my body would necessarily be to doubt my ability to move . . . I find my power of movement limited by the resistance of objects around me. This felt resistance to my will defines for me, in conjunction with my perceptions, my own situation as an object among other objects.[8]

And not only is this in fact so: according to Professor Hampshire, it could not conceivably be otherwise. It is only because persons are themselves physical objects with a situation in space and time and with a power of movement which brings them into contact with other physical objects, including other persons with whom they can communicate, that they can even form the idea of an objective world.

Professor Hampshire also thinks that the concept of human action provides the key to the problem of personal identity.

> We have no reason to seek for some criterion of personal identity that is distinct from the identity of our bodies as persisting physical objects. We find our intelligence and our will working, and expressing themselves in action, at a particular place and a particular time, and just these movements, or this voluntary stillness, are unmistakably mine, if they are my actions, animated by my intentions . . . I can only be said to have lost a sense of my own identity if I have lost all sense of where I am and what I am doing.[9]

It appears that the action here envisaged is always at least partly physical; so that it follows, in Professor Hampshire's view, that the notion of a disembodied person, and, therefore, the notion of personal survival in a disembodied state, is self-contradictory or meaningless. Mr. Strawson does not go quite so far. As we have seen, he thinks that there could not be an underived concept of a pure individual consciousness, but he sees no reason why such a concept should not have what he calls "a logically secondary existence." One can, therefore, intelligibly think of oneself as surviving one's

bodily death. For one can imagine oneself continuing to have experiences of various kinds, without having any power to make physical changes in the world, and without having any perception of a body which is related to these experiences in the way that one's living body is related to one's present experiences. I suppose it might be necessary to add the further condition that one's experiences should not in any way suggest that other people perceived such a body either. If these conditions were fulfilled, then one could legitimately think of oneself as surviving in a disembodied state. It is not suggested that this could actually happen, in the sense that it is causally possible for there to be experiences which are independent of a body, but only that the idea is intelligible. Having made this concession, Mr. Strawson goes on to remark that there are two essential features of this form of existence which may somewhat diminish its appeal. The first is that one would be entirely solitary; if there were other creatures in the same condition one would have no means of knowing it. The second is that one could retain one's sense of one's own identity only insofar as one preserved the memory of one's embodied existence; this might be eked out by taking a vicarious interest in the state of the world which one had left. In short one would exist, as it were on sufferance, as a former person. From this point of view the idea of there being persons, even of such an attenuated sort, who were not at any time embodied, is not intelligible.

*2*

I have made a detailed summary of Mr. Strawson's theory because it gives an account of persons which, if it were acceptable, would remove many of the difficulties of the mind-body problem. It seems to me, however, that the theory has serious difficulties of its own. The cardinal point is the attempt to stop short of physicalism on the one hand, and dispense with the argument from analogy on the other, by maintaining that our observations of the physical condition

and behavior of other persons, on the basis of which we attribute experiences to them, are logically adequate for this purpose. But what exactly is meant here by saying that a criterion is logically adequate? Not that the evidence entails the conclusion, for in that case we should not stop short of physicalism: if a statement about a person's experiences is to follow logically from a statement about physical events, it also must be construed as a statement about physical events. Not that the evidence provides sufficient empirical support for the conclusion, for then the reasoning is inductive: we are back with the argument from analogy. What is envisaged is something between the two, but what can this be? What other possibility remains?

That there can be a relation between statements which is not deductive and yet is in some sense logical is a view which Mr. Strawson is not alone in holding. It is maintained also by the followers of Wittgenstein, especially in connection with this problem, but they too fail to make it sufficiently clear what the relation is supposed to be. What Wittgenstein himself appears to have held is that it is only insofar as our so-called inner experiences have characteristic outward expressions that the statements which we make about them can have any meaning for us; and that so far as this goes it makes no difference whether one is referring to one's own experiences or to those of other people; it is in this sense that he denied the possibility of a private language. I do not think that he was right on this point, as I have argued elsewhere, but even if he were right, even if it is only through their having physical manifestations that our experiences are communicable, even to ourselves, the relation between the statements which refer to these experiences and those which refer to their outward expressions remains obscure. We are not allowed to say that the experiences are identical with their outward expressions; and yet we are not allowed to say that they are logically distinct. This would seem to indicate that there is a relation

of one-way entailment; but the entailment cannot run from the manifestations of the experiences to the experiences themselves, for the manifestations may be deceptive; and if it goes in the reverse direction, then in talking about our experiences we must be talking about their outward expressions and something else besides. But then the question arises what is this something else besides; and to this we are not given any answer.

But may not the reason why we get no answer be that the question itself is wrongly framed? If we begin by assuming a dichotomy between experiences on the one side and physical states or processes on the other, we shall surely end in the unhappy position of having to find some inductive ground for bridging the gap between them. But is this not just the assumption that Wittgenstein was trying to discredit? His followers will claim that our notion of a "pure experience" is utterly obscure to them. We talk of the experience of feeling pain, but do we really understand what it would mean to be in pain without having at least the tendency to display some physical reaction? Can we significantly divorce our thoughts and our emotions from their characteristic expressions in action or in speech? And when it comes to the way in which one observes the behavior of another person, it is surely quite wrong to treat this as an ordinary instance of the observation of physical events. Human behavior does not present itself to us as a physical process from which we have to make a dubious inference to the thoughts and feelings and purposes which lie "behind" it. It is itself expressive of these thoughts and feelings and purposes; and this is how we actually see it. From the outset we observe it *as* human behavior with all that this implies.

I do not question the facts on which this argument is based. No doubt we attach a significance to human behavior which we do not attach to the movements of inanimate things; there is, indeed, a sense in which one can simply observe what an-

other person is thinking or feeling. But however natural this process may be, it is still a process of interpretation; there is a distinction to be drawn between the sign and what it signifies. However intimate the relation between our "inner" states and their "outward" expressions, it is surely a relation between distinguishable terms. Indeed this is already implied by saying that the outward experience is a criterion for the existence of the inner state. But then we are entitled to ask what sort of criterion it is and what can be meant by the claim that it is logically adequate.

Mr. Strawson himself has tried to illustrate what he means by logical adequacy by appealing to another example.

> If one is playing a game of cards, the distinctive markings of a certain card constitute a logically adequate criterion for calling it, say, the Queen of Hearts; but in calling it this, in the context of the game, one is ascribing to it properties over and above the possession of these markings. The predicate gets its meaning from the whole structure of the game.[10]

In the same way, he suggests, the physical criteria which are held to be logically adequate for the ascription to persons other than oneself of predicates which imply the presence of consciousness do not exhaust the meaning of these predicates. The predicates get their full meaning from the structure of the language.

But what does this analogy come to? It is quite true that we recognize a card, such as the Queen of Hearts, by its markings; and it is also true that, in most contexts, when we identify a card as the Queen of Hearts we are saying not merely that it has a certain characteristic appearance, but also that it stands in certain relations to other cards, that it occupies a certain position in one of the series that makes up a suit. Even to say of it that it is a card implies that it is meant to figure in a game. This is one of the numerous cases in

which the applicability of a term to a given object depends, in part at least, upon the object's function. All the same it is always a contingent fact that a thing with such and such a characteristic appearance fulfills the function that it does. If the observation that it has certain markings is an adequate criterion for a particular card's being the Queen of Hearts, the reason is that the role which is played by the Queen of Hearts in various card games is commonly bestowed on cards of that design. But this is an inductive generalization; there is no logical connection between the fact that a card looks as it does and the fact that if it is used in a game it is allotted certain powers, for example that of outranking other cards whose markings are in some respects like and in other respects unlike its own. That a card has the look of the Queen of Hearts does not itself guarantee that it would be suitable to play the role of the Queen of Hearts in any particular sort of game, or indeed that it is fitted for a part in any game at all. The correlation of its appearance with its function is a matter of convention; but it is an empirical fact that this convention holds.

This may, however, be mistaking the point. Perhaps the point is that, given the appropriate conventions, it is not a matter for empirical discovery that a card with such and such markings plays such and such a role. In the context of a game of bridge, to identify a card as the Queen of Hearts is to identify it as a card which outranks the Knave of Hearts and is outranked by the King. The reason why the appearance of the card is a logically adequate criterion for its function is that the connection between them is established by the conventions which allot to cards of various designs their respective powers in the game.

But if this is the point of the analogy, it does not achieve its end. For the connection between a mental occurrence and its bodily expression is just not on a par with that which is conventionally established between the appearance and function

of a token in a game. To identify a piece in a game of chess as a bishop and to count anything other than a diagonal move with it as a move in the game would be a contradiction; the rules of chess being what they are, the identification of a piece can be construed as carrying with it the delimitation of its powers. On the other hand, there would be no contradiction in identifying a man's grimace as one which was characteristic of a man in pain, and yet denying that he felt any pain at all. The grimace and the feeling are logically separable in a way that, given the appropriate conventions, the appearance and function of a token in a game are not.

It has been suggested to me by Professor Alston that the analogy may hold in a weaker form. So long as we confine our attention to particular instances, we shall not find anything more than an empirical connection between a mental occurrence and its bodily manifestation. It will always be logically possible in any given case that either should exist without the other. And if this is possible in any given case, it is natural to infer that it must be possible in all cases. The view suggested to me by Professor Alston, which he thinks may also have been held by Wittgenstein, is that this inference is incorrect. Though the liaison between the characteristic outward expression of an inner state and the inner state in question may fail in any particular instance, it is not logically possible that it should fail in all instances, or even in any high proportion of them. So the reason why behavioral criteria can be said to be logically adequate is that even though they are not infallible, their overall success is logically guaranteed.

The source of this guarantee is supposed to lie in the fact that it is only through their being associated with certain outward expressions that we are able to talk significantly about our inner experiences. As has already been noted, this is Wittgenstein's ground for denying the possibility of a private language. We are taught the use of a word like "pain" in

contexts in which the feeling for which it stands is outwardly manifested in some characteristic way; and the result is that this association is retained as part of the meaning which the word has for us. The association is not so close as to exclude the possibility of anyone's ever feeling pain without displaying it, or of anyone's ever displaying signs of pain, without actually feeling it, but it is close enough to make it a logical certainty that such cases are the exception and not the rule.

One of the attractions of this theory is that it bars the skeptical approach to the problem of one's knowledge of other minds. For if we have the a priori assurance that the passage from outward manifestation to inner state is generally secure, we need no further justification for trusting it in any given instance; the onus then falls upon the skeptic to show that in these special circumstances it is not to be relied on. The question is, however, whether we are entitled to this assurance; and here I am still disposed to think that the skeptic can maintain his ground. For even if one grants the premise that we should not in practice be able to acquire an understanding of words which refer to inner states or processes, unless these inner states were outwardly detectable, it does not seem to follow that once our understanding of these words has been acquired, we cannot divorce them from their original associations. Indeed it is admitted that we can do so in particular instances; and it is not clear what should prevent us from doing so in all. If the suggestion is that we should then be landed in a contradiction, I can only remark that I do not yet see where the contradiction lies.

Whatever difficulties there may be in supposing that our mental states could in general be dissociated from their characteristic physical manifestations, the admission they can be so dissociated even in a few particular instances is fatal to the strict theory of physicalism. For if to speak about a so-called mental event were always logically equivalent to speaking about its physical manifestations, it would not be possible

even in a single instance that one should exist without the other. Physicalism of this type again has the merit of removing any difficulty about one's knowledge of other minds; but it achieves this by too desperate a measure. The decisive objection to it was put most succinctly by Ogden and Richards in their *Meaning of Meaning:* one cannot apply it to oneself except at the cost of feigning anesthesia.[11] Not only do we not have to observe our own behavior, or take account of our physical condition, in order to know what experiences we are having, but in many cases at least the occurrence of the experience appears to be logically consistent with the absence of its customary outward expression, or indeed of any outward expression at all. The exceptions are those cases where the words which we use to refer to the experience already associate it with some pattern of behavior: this applies especially to words which stand for emotions, perhaps also to statements of intention. But even here it would seem that if we thought it useful, we could cut away the references to behavior, and thereby obtain statements which were understood to refer to the experience alone. Even if it were true, as Wittgenstein seems to have thought, that our ability to talk significantly about our "inner" experiences depends upon their having characteristic outward expressions, it still would not follow that these outward expressions could not be deceptive; not just in the sense that they failed to cohere with other physical manifestations, but in the sense that the person in question simply did not have the experience which they led us to attribute to him.

It has been suggested that even if statements about experiences are not logically equivalent to statements of a purely physicalistic character, it may still be the case that experiences are identical with physical events. To speak of the brightness of the Morning Star is not logically equivalent to speaking of the brightness of the Evening Star; the sense of these two expressions is not the same; nevertheless the fact is that the

Evening Star and the Morning Star are identical. So also lightning is in fact a discharge of electricity, though the terms are not logically equivalent. In the same way, it is argued that even though the reports which we make of our experiences do not entail any descriptions of the condition of our brains, it does not follow that the two are not to be identified. It may be that all that actually takes place in this connection is a physical process in the brain; and that when it is one's own body which is in question one is able to apprehend this physical process in the form of an experience.

The difficulty here is to see what can be meant by saying that our experiences are not merely caused by physical occurrences which take place in our brains, but are literally identical with them. How could this claim be tested? What kind of experiment would establish that they were or were not identical? In the case of the Morning and the Evening Star we have a criterion of identity; there is empirical evidence that a star which is observed to shine at one period of the day is spatio-temporally continuous with one which is observed to shine at a different period; but this criterion is not applicable here. The case of lightning is more promising; for part of our reason for identifying the flash of lightning with an electrical discharge is that the electrical discharge produces it; even so what is more important is that they occur at the same time and place. But although there are philosophers who have taken this view, it does seem rather strange to hold that our experiences are literally located in our brains. This is not, as in the other case, a spatial coincidence which can be empirically discovered; it is rather that if one makes the assumption that experiences must be somewhere in physical space, the brain seems the obvious place in which to put them. But of course the dualist's answer to this would be that they are not in physical space at all.

The most that can be empirically established is that our experiences are causally dependent upon the condition of our

brains. To go beyond this and maintain that what appears to us as a correlation of the mental with the physical is really an identity, is simply to take a decision not to regard the mental correlates as entities in their own right. But even if this decision were accepted, it would not dispose of all our problems; for the identification of experiences with events in the brain is based on the acceptance of psychological laws. And how do we know that these laws themselves are valid? A possible answer is that I can find out the connection in my own case, whether or not I choose to regard myself as "owning" the experiences with which my body is associated; and that I can then infer that what has been found to hold for this body holds for other bodies also. But this brings us back once more to the argument from analogy.

*3*

An essential feature of the argument from analogy is that the justification, as distinct from the cause, of my ascribing experiences to others must issue from the premise that I have experiences myself. The assumption is that I can come to think of myself as having both a body and a mind without having to raise the question whether or not I am unique in this respect; then, knowing myself to be a person in this sense, I can go on to consider what grounds I have for believing that others are so too. But this is just the assumption that Mr. Strawson wishes to deny. He maintains, as we have seen, that it is a necessary condition of one's ascribing states of consciousness to oneself that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself; and from this he infers that any attempt to justify the belief that there are other persons by relying on the premise that one knows oneself to be a person would be circular; the premise would already assume what the argument is supposed to prove.

I shall now try to show that this inference is not correct. It is, indeed, a mark of a general concept, such as the concept

of being a person, that no limitation is placed upon the number of individuals to which it can apply. It is, therefore, true that I could not think of myself as satisfying the conditions of being a person unless I admitted the possibility that others satisfied them too. But all that this excludes is the view, which has indeed been held by some philosophers, that it is meaningless to ascribe states of consciousness to anything but oneself; it is perfectly compatible with the view that one does not know, or even with the view that one can not know that such an ascription is ever true in fact. But if I can know that I am a conscious subject without knowing that there are any others, there need be no circularity in an argument which proceeds from the premise that I have experiences and arrives at the conclusion that others have them also.

To this it may be objected that when it is said that I must be prepared to ascribe states of consciousness to others in order to be able to ascribe them to myself, what is meant is not just that I must be willing to regard the existence of other conscious beings as something which is abstractly possible, but rather that I must be disposed to treat it as established by certain observations which I actually make. In other words, I must already believe that I can be justified in ascribing experiences to others before I can significantly ascribe them to myself, and in that case an attempt to base this justification on my knowledge of my own experiences will be circular.

But the answer to this is that there is a difference between my believing that I am justified in accepting a given proposition and my really being so; and not only that, but a belief in a proposition may be justified without the proposition's being true. So even if I could not think of myself as a person unless I also thought that I had reasons to think the same of others, I could still consistently raise the question whether these reasons really did the work that was required of them; and even if I were able to decide that they were good reasons, I

still should not be bound to hold that they were conclusive. It seems to me, therefore, that Mr. Strawson's argument needs a stronger premise than the one that he states. It must hold it to be a necessary condition of one's ascribing states of consciousness to oneself not merely that one should ascribe them, or be prepared to ascribe them, to others, but that one should be sure of doing so successfully. It must, in other words, exclude the possibility of one's being invariably mistaken.

This premise would establish Mr. Strawson's case if it were true. My objection to it is that it is false. I do not mean by this that one is invariably mistaken in ascribing states of consciousness to others: if this were put forward as a general proposition implying that everyone was so mistaken, it would be self-contradictory; and I certainly do not in fact believe that I am the only conscious subject in the world. What I mean is that if there were someone who was invariably mistaken in ascribing states of consciousness to others, whether because there were no other persons in the world or merely because he never encountered any, this would not necessarily prevent him from being able to ascribe them to himself. Since this degree of solitude has never been attained by anyone who has acquired the use of language, I cannot prove my point by citing any actual case. The only way I can substantiate it is by constructing a rather artificial example.

Imagine a child who, for reasons which need not trouble us, is kept from having any contact, at least throughout his formative years, with any other human being. He is fed by mechanical means, and confined to a nursery which contains, in addition to the usual furniture, a number of automata. These automata, some of which roughly imitate the appearance of human bodies, are so constructed that they respond to his actions in certain limited ways; cry out, for example, when he hits them, or retaliate by striking out at him; they can be made to nod or shake their heads in answer to his questions, and also to utter certain simple sentences; these utter-

ances may be triggered off by visual or auditory signals. For his instruction in the use of language, and in other forms of behavior, the child depends upon a voice which addresses him through a loudspeaker. If it be objected that even though the child never sees the owner of the voice, it still introduces a human element, we can suppose that the sounds which the child hears are not spoken by any person, but transmitted, perhaps from written messages, by a machine. The voice teaches the child the names of the various types of objects in the room; it formulates sentences for him to imitate, which are sometimes also responded to by the automata. It teaches him his own name, the use of pronouns and demonstratives, and the use of words which describe his "inner" states. In very much the same way as children normally do learn these things, he learns to say when he is hungry or satisfied, happy or in pain; he is coached, as other children are, to distinguish what he sees from what he imagines, or from what he remembers; and among his memories to distinguish those that are memories of dreams. Part of the method by which this is achieved is a stressing of the similarity between himself and the automata; the voice always speaks of them as though they too were conscious, and he finds that the attribution of consciousness to them, but not to the "inanimate" objects in the room, corresponds to differences in their behavior. In this way he learns how to apply the concept of a person: and he satisfies the condition of being ready to apply it to other things beside himself.

I said that this was an artificial example, but it is not excessively fanciful. The idea of such an experiment may be morally repugnant, but it would not be very hard to stage, and I see no reason to suppose that it could not lead to the result which I have described. But if this is so, the consequences are important. The example shows not only that one might be able to ascribe experiences to oneself, while being invariably mistaken in ascribing them to others, but also

that the criteria which are taken to be logically adequate for ascribing experiences to others may determine no more than that some locution is correct, that in such and such conditions this is the proper thing to say; it does not necessarily follow that what is said is true. Nor can this conclusion be escaped by saying that the child in my example would not possess our concept of a person, that the automata which he had been taught to regard as conscious subjects really would be persons in his sense of the term. There is no warrant for assuming that his concept of a person could not be the same as ours. He applies it to himself on the basis of his experiences and he applies it to the automata on the basis of their behavior, which if they were very skillfully constructed, might not appear very different, within their limited field of operation, from that of human beings. It is not that he has a different concept of what it is to be conscious, or that he applies the concept incorrectly, but that he just happens to be in a situation where the things which he has every reason for thinking to be conscious are really not so. If he were an infant philosopher, he might begin to wonder whether his companions really did have experiences in the way that he did and infer that they did from their resemblance to himself. Or perhaps if he were struck by some stereotyped quality in their behavior he would rightly conclude that they did not. Whichever conclusion he came to, his skepticism would not be senseless. How could it be if it were actually justified?

The reason for saying that a doubt of this kind would be senseless, even in this artificial instance, is that it could never be laid to rest. There is nothing in this child's world, as we have described it, which would allow him to find out that his companions were, or were not, automata. If he were released from his nursery and allowed to mix with other human beings, he would have a standard of comparison. He would be able to observe how differently these human beings spoke and acted from the things which he had been brought up to

regard as persons; if he discovered how the automata were constructed, he would have a further reason for placing them in a different category from himself. But would this mean more than that he was requiring a stronger basis for his argument from analogy? And might he not still doubt whether it was strong enough? But then what would make it strong enough? If no behavioral criteria are logically sufficient, there must always be room for doubt. But what is the significance of a doubt which could never be allayed?

I do not know the answer to this question. The difficulty is not just that the argument from analogy, at least in the form in which it is usually presented, does not seem very powerful; it is rather that any inductive argument allows for its conclusion to be false. So if my belief that other beings are conscious can be defended only as an inference from their behavior, it is at least possible that I am the only person in the world. The short answer to this is that I know that I am not; but this still leaves the problem how I can have the right to be so sure. If physicalism were true, this problem would give no trouble, but I do not see how the obvious objections to physicalism are to be overcome. I am, therefore, inclined to agree with Mr. Strawson that we ought to try to find a middle way. For the reasons which I have given, I do not think that he has yet succeeded: among other things, his notion of logical adequacy is rather too nebulous to bear the weight which he puts upon it. Nevertheless, I think that the development of this notion, along the lines suggested by Professor Alston, may at least be along the road in which the solution of this problem lies.

Though I am by no means fully satisfied with the argument from analogy, I think it can be defended against the objection that the only experiences which it can give us any reason to believe in are experiences of one's own. This objection rests on the premise that if the behavior of others did not provide me with a logically adequate criterion for ascribing experi-

ences to them, the very idea of their having experiences could have no meaning for me. The only facts in my possession would be that when my body was stimulated in certain ways, certain experiences occurred, and that when bodies similar to mine were stimulated in similar ways no such experiences occurred. But if this were all I had to go on, then, it is suggested, I could not even form the hypothesis of there being experiences which were enjoyed by persons other than myself.

This objection has force when it is backed by the assumption that my successfully ascribing experiences to others is a precondition of my being able to ascribe them to myself; but when this assumption is discarded, all that remains for the objection to rely on is a strict interpretation of the verification principle. Its conclusion will follow from the ruling that I cannot attach sense to any statement which I could never be in a position to verify directly. But this in itself appears to me now to be an objection against maintaining the verification principle in such a stringent form. If no more is required than that the statement be indirectly testable, then the argument from analogy will at least not be excluded at the outset. Whether its conclusion satisfies this condition of significance will in fact depend upon the question whether it is a legitimate form of inductive argument. There is indeed a special difficulty for those who think that one can not conceive of physical objects except as logical constructions out of one's own private sense-data; for it may well be asked how I could possibly suppose that a logical construction out of my sense-data was endowed with a private world of its own. On the other hand, if I am entitled to assume that what I perceive, in some cases at least, are public objects which exist independently of myself, there seems to be no good reason why I should not be able to form the hypothesis that certain of these physical bodies have experiences connected with them in the way that experiences have been found by me to be connected with particular states of my own body. The objec-

tion that I must in that case already be conceiving of the physical bodies as the bodies of other persons does not seem to hold; the identification of such a body as a particular physical object will, as I hope to show, itself be sufficient for the numerical identification of the experiences which are causally dependent on it; further than that, the experiences themselves need be identified only in a qualitative fashion, as experiences of a certain kind.

If we are able to dispense with the notion of a subject, other than the body, to which both states of consciousness and physical properties are attributed, there will be no reason why we should especially direct our attention to predicates which are concerned with human action. The importance which was attached to these predicates was that they were supposed to provide us with an instance in which we had knowledge not based on observation and inference of something other than our own experiences. Thus our immediate knowledge of our own bodily movements would supply a precedent for taking our knowledge of the experiences of other persons to be noninferential. But, if, as I have tried to show, our knowledge of the experiences of others is inferential, the need for this precedent will not arise. In any case, it seems to me very doubtful whether the precedent itself is valid. It is indeed true that we are made aware of our bodily movements by kinesthetic sensation, and it is also true that we usually do what we intend to do, especially when the intention is one that can be immediately carried out. Nevertheless our kinesthetic sensations can be delusive; they are not sufficient conditions of the physical movements with which we associate them; neither is one's consciousness of an intention a guarantee that it will be fulfilled. In both instances the exceptions are rare enough for us to overlook the distinction between the sign and what it signifies; we speak of feeling a movement or of doing things intentionally, as though the sensation were inseparable from the movement,

and the intention from the action which fulfills it; but the distinction is there none the less.

Neither does it seem that the analysis of predicates, which are concerned with human action, will provide us with the answer to the problem of self-identity. Professor Hampshire may well be right in claiming that even in our infancy we are not simply the passive recipients of sensory impressions; it may be true that one comes to think of oneself as an object among other objects through feeling the resistance which these other objects oppose to one's will. But this fact, if it be a fact, about the way in which a person acquires the idea of his own identity does not tell us in what this identity consists.

*4*

If the thesis of physicalism is to be rejected, it does not follow that we are forced back into the position of Descartes. It is possible to hold that states of consciousness are distinct both from their physical causes and from their physical manifestations, without being obliged to hold that there is a mental substance to which they are ascribable. Indeed, if this is what Hume was seeking when he tried to distinguish himself from his perceptions, it is not merely a contingent fact that he was unable to find it; there is no possible way in which such a substance could be identified.

It might seem, however, that we were then committed to Hume's own position. If experiences can be logically distinguished from physical events, then it might appear to follow that one can at least conceive of their existing on their own. Having rejected the notion of a mental substance, we shall have to look upon the self as "a bundle of perceptions," in Hume's terminology; and it will be a contingent fact that separate bundles are associated with particular bodies. If we take the short step from Hume to John Stuart Mill and regard these bodies themselves as "permanent possibilities of sensation," it will be a contingent fact that some particular

"perception" is both an element of a collection which constitutes a certain body and an element of a collection which constitutes an individual self.

But the answer to this is that all that we have so far allowed is that a statement which ascribes an experience to some person need not be equivalent to any statement which refers to that person's physical condition or behavior. And from this it does not follow that it is even logically possible for states of consciousness to exist independently of any physical body. The reason why it does not follow is that it may not make sense to talk of states of consciousness except as the experiences of some conscious subject; and that it may well be that this conscious subject cannot be identified except by reference to his body.

In favor of this view, it may be argued that the alternative of equating the conscious subject with the series of his experiences leaves us without any explanation of the nature of personal identity. Not only is it not clear how the individual experiences are to be identified, but there appears to be no principle according to which they can be grouped together; there is no answer to the question what makes two experiences, which are separated in time, the experiences of the same self. The most promising suggestion is that the bundles are tied together by means of memory; but this meets with serious difficulties. In the first place, it is exposed to the charge of circularity; for it is plausible to argue that remembering an experience already implies thinking of it as an experience of one's own: and even if this charge can be met, it is clear on other grounds that memory alone will not suffice. For not every experience can be remembered; otherwise each piece of remembering, which is itself an experience, would have to be remembered, and each remembering of a remembering and so *ad infinitum:* how then is it to be determined that two memory experiences which occur at different times

are members of the same bundle? The only answer which suggests itself is that one of them accompanies an experience of which the other is directly or indirectly a memory: the relation may be said to be indirect when the second memory is a memory of some experience accompanying a memory which is either directly, or at one or more such removes, a memory of the experience in question. But what is this relation of accompanying? When dealing with the question in my *Problem of Knowledge* I said that it might be taken to be "the relation that holds between two items of experience if and only if they are parts of the same total experience at any given moment."[12] But this does involve us in a circle, for

> what is meant here by a total experience is just the experience of one and the same person. We can hold that the relation between its parts is *sui generis,* but then we can also hold that the relation between the successive experiences of the same person is *sui generis;* and in that case we do not need to bring in memory at all.[13]

In the *Problem of Knowledge* I did indeed fall back upon a solution of this kind: I spoke of a "relation of which, perhaps, nothing more illuminating can be said than that it is the relation that holds between experiences when they are constituents of the same consciousness."[14] But to be driven to postulating an unanalyzable relation is, in this as in other instances, simply to abandon the problem, not to solve it.

One objection to making personal identity depend upon the identity of the body is that it rules out even the logical possibility of a person's existing in a disembodied state. On any view, the evidence that states of consciousness are causally dependent upon physical processes is strong enough to make it extremely improbable that any person ever does succeed in doing this; but the idea of its happening might at least appear to be intelligible. Surely one can imagine oneself continuing to have experiences which are very like the ex-

periences which one has at present, except that they include none of the perceptions which establish the existence of one's own body; and if one can imagine that this could happen to oneself, one should also be able to imagine that it could happen to others. This is a legitimate form of argument, but in the present instance it may be deceptive. For if this picture of oneself in a future disembodied state is intelligible in its own right, why should it have to contain any link with one's present form of existence? Is it not also imaginable that one should lose all memory of one's embodied self? And could we not go even further? If it is conceivable that one should exist at some time without having any experiences which were indicative of one's having a body, why should it not be conceivable that this should be so at all times? Could it not be imagined that the whole of one's existence was passed in a disembodied state? The objection to these flights of fancy is that there would then be no means by which one could be identified. But does this not equally apply to the idea of a person's surviving the destruction of his body? The experiences which might then ensue may themselves be imaginable; but it would seem that in crediting them with an owner we are making a tacit reference to the body which is supposed to have been forsaken. Mr. Strawson's idea appears to be that one could retain one's identity through having memories of one's former life. But here he seems to have forgotten that something is needed to secure one's continued existence as a person; and for this we have seen that memory will not suffice.

I do not present these arguments as being in any way conclusive; I am, however, inclined to think that personal identity depends upon the identity of the body, and that a person's ownership of states of consciousness consists in their standing in a special causal relation to the body by which he is identified. I am not maintaining of course that this is how one actually becomes aware of one's own experiences, but

only that the fact that they are one's own, or rather the fact that they are the experiences of the person that one is, depends upon their being connected with this particular body. This amounts in effect to adopting what Mr. Strawson calls "the no-ownership doctrine of the self." We must, therefore, try to rebut his charge that this theory is internally incoherent.

His argument, as we have seen, is that the theory requires it to be a contingent proposition that all my experiences are causally dependent upon the state of my body; but if my experiences are identified as mine only in virtue of their dependence on this body, then the proposition that all my experiences are causally dependent on the state of my body must be analytic; and so the theory is committed to a contradiction. I think, however, that these propositions can be reformulated in a way which shows that the contradiction does not really arise. The contingent proposition is that if my body is in such and such a state then an experience of such and such a kind results; the analytic proposition is that if an experience is causally dependent in this way on the state of my body, then the experience is mine. But now it is obvious that these propositions are distinct; so that there is no inconsistency in holding that one is contingent and the other not. There would indeed be a vicious circle if the experiences had first to be identified as mine before it was discovered that they were dependent on my body, but this is not the case. The position is that a person can be identified by his body; this body can be identified by its physical properties and spatio-temporal location; as a contingent fact there are certain experiences which are causally connected with it; and these particular experiences can then be identified as the experiences of the person whose body it is. There is nothing inconsistent in this.

What may have misled Mr. Strawson is the picture of a heap of experiences which have to be assigned to their re-

spective owners. For this makes it natural to ask which of these experiences are dependent upon a given body, and to believe that one is raising a question of fact; it will then seem to be an objection to our theory that it only allows the trivial answer that the experiences which are dependent on this body are those which are dependent on this body. But it is the picture that is at fault; and the question to which it leads is illegitimate. We cannot ask *which* experiences are dependent upon a particular body, and are, therefore, to be assigned to such and such a person, because this is to assume improperly that the experiences have been independently identified. The question which we can ask is *what* experiences are dependent upon a particular body, if this is just a way of asking what experiences the person who is identified by the body is having at the relevant time. It is analytic that if the experiences are connected with his body, they are his experiences; but of course it is not analytic that experiences of one sort rather than another are at any given time connected with his body. Neither does our theory require that it should be.

Our difficulties, however, are not yet at an end. Merely to say that a person's experiences are causally dependent on the state of his body is to speak too vaguely. The nature of the causal relation which is to do the work of assigning experiences to persons needs to be precisely specified. If we are to do justice to the assumptions that are ordinarily made about the way in which experiences are distributed, this relation must operate in such a way that any individual experience is linked to one and only one human body. The problem is to find a causal relation which not only fulfills this task but is also such that its existence can plausibly be taken to follow from every statement in which a person is credited with some experience.

A simple answer would be to regard the relation as being that of causal sufficiency. An experience would then be said

to be mine if and only if some state of my body were causally sufficient for its occurrence independently of the state of any other body; this would exclude the experiences of other people on the assumption that I cannot act upon another person except by producing some physical change in him which is then itself sufficient for the occurrence of his experience. But apart from its needing this questionable assumption, the objection to this answer is that it relies on a causal relation, the existence of which we are hardly entitled just to take for granted. It is by no means universally agreed that all our experiences are physically determined, in so strong a sense as this would imply; it has even been maintained, in the interests of free will, that there are experiences which have no sufficient conditions at all. Now this may very well be wrong; there is certainly no means of showing that any experience lacks a sufficient condition, and it may in fact be true of every experience that its sufficient condition consists in some state of its owner's body. But the point is that even if this hypothesis is true, it can significantly be questioned. It is not at all plausible to hold that its truth is logically implied by every statement in which a person is credited with some experience. To say that the experience is not physically determined may be false, but it does not appear to be self-contradictory.

I think, therefore, that the most that we can hope to maintain is that an experience belongs to a given person in virtue of the fact that some state of that person's body is a necessary condition of its occurrence. The justification for this would be first that experiences are individuated only by reference to the persons who have them, and secondly that persons are identified only by reference to their bodies. If these premises are admitted, it follows that no experience of mine would have occurred unless my body had existed; more particularly, it follows that the existence of my body will be implied by any statement in which an experience is attributed to me. But this does not settle the argument. For even if it were

granted that an experience could not be mine unless it stood in this relation of causal dependence to my body, one might still maintain that something more than this was required to identify the experience as mine.

A strong argument in favor of this view is that if the relation of dependence is to be merely one of causal necessity, then every one of my experiences is dependent upon the existence of bodies other than my own. This follows simply from the fact that I must have had ancestors: since the existence of their bodies is a necessary causal condition of the existence of my body it is also a necessary causal condition of the existence of my experience. Moreover, apart from this general condition which applies indifferently to all my experiences, a great many of my experiences owe their special character in part to the behavior of other persons. It is very often the case that I should not be having the particular experience that I am having unless some other persons had spoken or acted in some particular way, or been in some particular state; and in all such instances the existence of these person's bodies will also be a necessary condition for the occurrence of my experience. It would appear, therefore, that while the relation of causal sufficiency is too strong for our purposes, the relation of causal necessity is too weak. It does not fulfill the task of assigning each experience to one and only one body.

The only way that I can see of overcoming this difficulty is to make a distinction between mediate and immediate necessity. Let us say that the existence of an event x is mediately necessary for the existence of another event y if and only if there is some event z such that x is a necessary condition for the existence of z and z is a necessary condition for the existence of y; and let us say that the existence of x is immediately necessary for the existence of y if and only if it is necessary for the existence of y, but not mediately so. Then we may claim that the causal relation which links a person's experi-

ence to his and only his body is that of immediate necessity. What makes a given experience mine is the fact that the existence of some state of my body is an immediately necessary condition of the occurrence of the experience and that no state of any other body is so.

Let us now see whether this criterion gives an adequate result. If our reasoning has been correct, the first of its requirements presents no difficulty. Not only is the existence of my body a necessary condition for the existence of any of my experiences, but it also seems clear that it is immediately necessary. There is no other factor that intervenes between my body and this set of experiences which are dependent on it: indeed it is hard even to imagine in this case what such an intervening factor could be. The question is whether there is any difficulty about the second requirement. Does it safeguard us from having to assign experiences to what would ordinarily be regarded as the wrong owners?

With some misgiving, for reasons which we shall come to, I am prepared to maintain that it does. The fact that the physical existence of my ancestors is also a necessary condition of my having any experiences now presents no problem. For clearly this is a case of mediate necessity. The existence of my own body is an intervening factor. It is also an intervening factor in at least the overwhelming majority of cases in which the character of my experiences depends upon the state or behavior of another person. For in the normal course of things the only way in which another person can act upon me is by affecting my perceptions. If I am to be in any way influenced by him I must observe him, or observe some effect of what he has done. But in that case the existence or state of his body can at best be a mediately necessary condition of my having the experience which depends upon it. It is mediated by the occurrence of my perception, and so by a bodily state of my own.

We could leave the matter there were it not that we must

allow at least for the possibility of para-normal experiences. Thus, there are alleged to be cases in which one person acquires information about the mental or physical condition of another without having to rely upon any form of sense perception. Now one may be skeptical about the authenticity of these reported cases of telepathy, though some of them at least appear to be very well-attested; one may accept them as authentic, but still believe that they can somehow be explained in physical terms. This would be in line with the assumption that it is impossible to act upon another person except by causing him to undergo some physical change. Nevertheless, however little we may like the idea of accepting telepathy at its face value, that is, as a form of communication between persons which does not operate by physical means, I do not think that we are entitled to exclude it a priori. It may be a far-fetched notion, but it does not appear to self-contradictory. But if we allow this, then the adequacy of our criterion for assigning experiences to persons is put in question. For suppose that someone communicates a thought to me in this telepathic fashion. In that case, his bodily state will be a necessary condition for his having the thought which he communicates; it will, therefore, also be a necessary condition of my having the experience which results. But *ex hypothesi,* the state of my own body is not in this instance an intervening factor. It is of course itself a necessary condition of my having the experience, but an independent one. It might, therefore, be thought that according to our criterion we should be obliged in a case of this kind to deny me the ownership of the experience.

This would, however, be a mistake. For what is required of our criterion is that no experience of mine shall have for an immediately necessary condition the state of any other body but my own. And this requirement is satisfied even in the case which we have just been envisaging. If we allow telepathy of this kind to be possible, then we are indeed al-

lowing that the state of another person's mind may be an immediately necessary condition of my having some experience; but the state of his body will still be only mediately necessary. It will give rise to my thought only through the medium of his. There is, therefore, still an intervening factor: not, as in this normal case, some state of my own body, but the other person's experience.

The only type of para-normal case which we could not so easily accommodate would be that in which my experience was dependent on the action of another person, without my being in telepathic communication with him and without my perceiving his action or any of its physical effects. Thus in an experiment on extra-sensory perception, in which one of the experimenters selects a card and the subject, sitting in another room, is required to guess what it is, it may be discovered that the subject scores significantly better when the experimenter touches the back of the card with his finger, even though the subject does not know what the experimenter is doing, and the experimenter himself does not know what the card is. If this were to happen, there would be a ground for saying that the experimenter's bodily movement was an immediately necessary condition of the subject's thinking as he did. But then we should be faced with the impossible consequence that, according to our criterion, the subject's thought was not ascribable to any single owner.

This example is troublesome, but not, I think, necessarily fatal to our theory. One way of meeting it would be to construe our criterion in such a way that the only type of bodily state which came within its scope would be an internal state. Such overt performances as the movement of a finger would not qualify. So certain experiences would be mine in virtue of the fact that such things as the condition of my brain and nerves were immediately necessary for their occurrence, and that they did not stand in precisely this relation to any other body but my own. Another defense would be simply to re-

fuse to admit our para-normal case as a counterexample. It is not possible, one might argue, that the experimenter should affect the subject's thought under these conditions simply by moving his hand. This type of action at a distance is unintelligible. There must be some intervening mechanism, even though we do not know what it is. But the trouble with this defense is that it already assumes that the subject's thought has been assigned to him. Without this assumption the complaint about action at a distance would have no basis as it stands. We could, however, attempt to modify it in such a way as to avoid this objection. We should have to maintain, as a general principle, that in a case where an experience would be manifested in a given body, if it were manifested at all, it was impossible that it should be causally dependent upon a different body, without the operation of some intervening factor. But how are we to decide which is the body in which the experience would be manifested? If we are reduced to identifying it as the body of the person who has the experience, then clearly we are back in our circle; and it does not seem certain that it could always be identified by other means. In view of these difficulties it seems preferable to adopt the course that I first suggested: that is, to try to deal with the awkward para-normal example by narrowing our interpretation of what is to count for our purposes as a bodily state.

A more far-reaching objection to this whole procedure is that we are introducing a recondite, if not dubious, theory into the analysis of statements which function at a much simpler level. The use of statements which ascribe experiences to persons is an everyday occurrence: one of the first things that a child learns is to employ and understand them. Can we seriously maintain that these statements incorporate such a sophisticated and unfamiliar notion as that of an immediately necessary condition?

The answer to this objection is that it is beside the point.

In attempting to analyze statements about persons, we are not proposing to discover what those who make such statements commonly have in mind. Our aim is rather to redescribe the facts to which the statements refer in such a way that their nature becomes clarified; and for this purpose there is no reason why we should not resort to technical terms. In the same way, it is not a fatal objection to a causal theory of perception that a child may learn to talk of the physical objects which he perceives before he has acquired the notion of cause, neither is it fatal to a phenomenalist theory that comparatively few people understand what is meant by a sense-datum. The only relevant question is whether these theories are adequate to the facts: that is, whether they correctly represent the truth-conditions of the statements which they serve to analyze. In the present instance, the way to refute our theory would be to find an example in which it clearly made sense to speak of a person's having some experience, even though the experience was not uniquely dependent on his body in the way that the theory requires. So long as no counterexample is forthcoming which the theory cannot be adapted to meet, we may regard it as provisionally acceptable. In a field in which so many theories have had to be discarded, I should not wish to claim more for it than that.

In claiming even so much, I am assuming that it has at least been shown that the theory is free of any vicious circularity. But this may still be questioned. The reason why it may be questioned is that the theory presupposes the existence of psycho-physical laws. Admittedly they are fairly modest laws: we are not assuming that every experience is physically determined; the physical factor is taken only to be necessary, not sufficient. Even so, it may be argued, these laws have had to be empirically discovered. And how could we ever have set about discovering them unless the experiences, which were found to be correlated with certain physical states, had themselves been independently identified? But

this means that the charge of circularity returns in full force. How is it finally to be met?

The only way that I can see of meeting it is to draw a distinction between the general proposition that every experience is causally dependent, in the required sense, upon a body and the more specific propositions which describe the different forms that this dependence takes. The general proposition must be held to be necessary, on the ground that causal dependence upon a body is an essential part of what we mean by an experience. On the other hand, the more specific propositions are contingent. The precise nature of the psychophysical laws which correlate experiences of various types with certain sorts of physical conditions remains a matter for empirical discovery.

In taking the more general proposition to be necessary, I am not implying that in order to know that I am having some experience, I have first to find out that it is dependent on my body. I do not need to find this out, any more than I need to find out that this body is my own. The identification of the body, which carries with it the numerical identification of the experience, is a problem for other people not for oneself. The reason for this is that in referring to myself at all I am presupposing my ownership of this body; in claiming an experience as mine, I imply that it is dependent on this body and not on any other. This does not mean of course that my body could not have been qualitatively different; we are concerned here only with numerical identity. It is not a necessary fact that my body has the physical attributes that it does, but given that this is the body by which I am identified, it is a necessary fact that *this* body is *mine*.

But now a further question arises. If my argument is correct, it is essential that a person be identified at any given time by reference to some body. But is it essential that he be identified at all times by reference to the same body? It would seem natural and convenient to hold that it is, but the

consequence of this would be that certain hypotheses which have been thought to be significant, even if highly improbable, would be ruled out a priori. We should, for example, be making it logically impossible that a person should be reincarnated, or that two persons should exchange bodies with one another. Yet however fanciful a story like Anstey's *Vice Versa* may be, it is not ordinarily thought to be self-contradictory.

What makes such fantasies appear legitimate is that there are subsidiary criteria of personal identity which may at least be conceived as running counter to the main criterion of physical continuity. These are the criteria of memory and continuity of character. Thus, in Anstey's story, the ground for saying that Mr. Bultitude has been translated into the body of his schoolboy son is that from a certain moment onwards the person who is identified by the schoolboy's body displays the mental characteristics which previously belonged not to the son but to the father, and that it is the father's and not the son's experiences that he ostensibly remembers. In such a case, we could insist on saying that the persons who were respectively identified by the two bodies remained the same as before but that they had mysteriously acquired each other's character and memories; it does, however, seem a more natural way of telling the story to say that the two persons have exchanged bodies. On the other hand, even if some one could convince us that he ostensibly remembered the experiences of a person who is long since dead, and even if this were backed by an apparent continuity of character, I think that we should prefer to say that he had somehow picked up the dead man's memories and dispositions rather than that he was the same person in another body; the idea of a person's leading a discontinuous existence in time as well as in space is just that much more fantastic. Nevertheless, I think that it would be open to us to admit the logical possibility of reincarnation merely by laying down the rule that if a person

who is physically identified as living at a later time does have the ostensible memories and character of a person who is physically identified as living at an earlier time, they are to be counted as one person and not two. For given that this condition is satisfied, the question of their numerical identity is a matter for decision and not a question of fact.

But even if the subsidiary criteria of personal identity could in these strange cases be allowed to override the primary physical criterion, they are still parasitical upon it. It is only because the different bodies provide us with subjects of reference that we can entertain these queer hypotheses at all. What we should in fact be doing in these cases would be to revert to a Humean theory in which a person's identity is made to depend upon relations between experiences, irrespective of the body with which they are associated. But we have seen that a theory of this kind is not tenable unless the experiences themselves can be identified; and I have argued that the only way in which they can be identified is through their association with a body, the body being that which supplies an immediately necessary condition of their occurrence. It may well be thought a defect in my position that it requires the existence of these psycho-physical relations to be assumed a priori. But if this is a defect it is one that I see no way to remedy.

# COMMENTS

Edmund L. Gettier III

*Wayne State University*

*1*

I CANNOT in this comment discuss all that Professor Ayer has to say in his paper. I shall restrict my comments to section 4 of the paper, where he states views of his own. In order to make discussion easier, I should like to begin by listing five propositions. Throughout these comments assume that the variable "S" ranges over persons, "B" over bodies, "E" over experiences, and "x" and "y" over the class of contingently existing objects.

a) If S identifies S′, then S identifies S′ by reference to the body of S′.
b) If S identifies an experience E, then S identifies E by reference to the person who has *E*.
c) E is an experience of S if and only if there exists some internal state of S's body that is an immediate necessary causal condition of the existence of E.
d) Then S asserts or judges that Fx S has identified x.
e) Then S identifies S′ by reference to B (there is a relation R such that
   i) S knows that there is one and only one that has R to B, and
   ii) S refers to S′ as the one and only one person that has R to B.)

I believe that it is clear from Ayer's paper that he holds propositions (a) through (c) to be true. I think that he would

want to hold (d) or something like (d), for one of the points of the paper seems to be that for any group of particulars, if we cannot identify particular members of that group, then we cannot make assertions or have beliefs about particular members of the group. Thus, if we cannot identify disembodied persons, then we cannot make assertions about particular disembodied persons. I do not know whether Ayer would accept proposition (e) or not. I shall not at this moment anyway attribute proposition (e) to him.

It is obvious that the expressions "identified" and "by reference to" are key expressions in the sentences used to state propositions (a) through (d). Clearly, also, if one does not understand these key expressions, then one would not know what propositions are being expressed by (a) through (d), and hence, what propositions Ayer is here holding to be true. Certainly these expressions need some explanation, possibly even definition, since they seem to occur as technical terms in the sentences used to express propositions (a) through (d). However, in his paper Ayer does not give any explanation of them at all! And in the absence of such explanations, it seems fair to say that Ayer has not presented us with one clear view to be discussed.

In order to make it quite clear that the terms "identified" and "by reference to" need explanation, I shall present *prima facie* counterexamples to proposition (a) and (b), and maintain that in the absence of some elucidation of these two terms it is impossible to tell whether these apparent counterexamples are or are not really counterexamples.

First consider proposition (a). According to it, if I identify Albert Einstein then I must have made some reference to the body of Albert Einstein. But suppose on some given occasion I refer to Einstein as the person who first thought of the theory of relativity. Clearly I can so refer to Einstein. And just as clearly, the definite description that I used contained within it no expression that refers to Einstein's body. Now,

have I, by so referring to Einstein, identified him? It is impossible to tell from what Ayer says in his paper. Perhaps I can successfully refer to a man without identifying him. Perhaps all successful references to a person are identifications. Perhaps a successful reference to a person is an identification only if there is another person who witnesses the reference and knows to whom the reference is made. Suppose that by referring to Einstein by means of this definite description I *did,* according to Ayer, identify Einstein. Have I done so by reference to Einstein's body? Clearly, my description contains no expressions referring to the body of Einstein. But does it or does it not follow from this that I have identified him without referring to his body? Again it is impossible to tell.

Now, consider proposition (b). According to it I can identify a particular experience E only by making some reference to the person who has experience E. But suppose I have a pain in my shoulder, and, referring to the pain in my shoulder, I say to myself, not to someone else, "That pain is getting worse." It seems that I can refer to my own pains in this way, provided I am talking to myself. But again, the expression "that pain" by which I referred to the pain in my shoulder, does not contain any expressions referring to myself, or for that matter to any person. Have I by referring to my pain with the expression "that pain" identified the pain? Might this not be a case of what Ayer calls "ostensive identification"? Or, again, is it required for a successful reference to a pain to be a case of identification, that there be a hearer who, upon hearing the reference, knows to what object the reference is made? Whether these two cases of a reference are or ar not counterexamples to propositions (a) and (b) cannot be determined by anything that I can find in Ayer's paper. I feel, therefore, that I don't know exactly what view he is asking us to consider.

*2*

In spite of the obscurity surrounding the expressions "identified" and "by reference to," it seems possible to argue

that if (a), (b), and (c) are true, then (a) and (b) are both false. I shall first argue that if (a), (b), and (c) are all true, then (b) by itself is false. I shall then argue that from what was established in deducing the falsity of (b) it follows also that (a) is false.

Let us assume that $B_1$ is a particular body that has been identified without any reference to a person or to any experiences. Let us say, as Ayer asserts is possible, that body $B_1$ has been identified by its place at a certain time. Let us assume also that a certain person $S_1$ has been identified by reference to $B_1$, and that $B_1$ is the body of $S_1$. All of this is said to be possible in (a) and (b). Now consider the following list of definite descriptions:

f) The experience E such that:
   i) E is P, and
   ii) E belongs to $S_1$
g) The experience E such that:
   i) E is P, and
   ii) There exists some internal state of the body of $S_1$ that is an ICNC (immediate causal necessary condition) of the existence of E.
h) The experience E such that:
   i) E is P, and
   ii) There exists some internal state of the body $B_1$ that is an ICNC of the existence of E, and
   iii) $B_1$ is the body of $S_1$
i) The experience E such that:
   i) E is P, and
   ii) There is an internal state of $B_1$ that is an ICNC of the existence of E.

It follows from (b), no matter what the words "identified by reference to" mean, that it is possible to identify an experience by reference to the person who has E. If Ayer holds that (b) is true, then it would seem that he would hold that an experience can be identified by means of description (f). The

first constituent of (f), that E is P, is inserted because for most people if they are having one experience, they are having many. Hence, it would be impossible to refer uniquely to one particular experience of a given person by saying merely that it is the experience had by that person. Some other properties of the experience must be mentioned. In this case, let us imagine that the property of being P is the property of being a thought about the number eight.

The second constituent of description (f), *i.e.*, that E belongs to $S_1$, is analyzed by Ayer in proposition (c) above. If we substitute the analysis of "E belongs to $S_1$" for this expression in description (f), the result is description (g). If Ayer holds that an experience can be identified by description (f), then it would seem that he must hold that the very same experience can be identified by description (g), since he holds that (c) states a logical equivalence.

Description (g) contains the expression "the body of $S_1$." We are assuming that a body $B_1$ has been identified, and that $B_1$ is the body of $S_1$. Hence, if we substitute the name "$B_1$" for the expression "the body of $S_1$" and then add a condition to the effect that $B_1$ is the body of $S_1$, we should get a description that uniquely refers to the same experience that is referred to by description (g). By making this substitution and addition we get description (h). It seems, then, that if Ayer holds (b) true, he would hold that it is possible to identify an experience by means of description (h).

But I think that it can be shown that if (h) uniquely refers to some experience, then (i) uniquely refers to an experience. We can see this by considering the transposition, i.e., if (i) does *not* uniquely refer to an experience, then (h) does *not* uniquely refer to an experience. Let us suppose that (i) fails to refer uniquely to an experience because there is more than one experience that satisfies this description. If description (i) fails in this way then there are at least two experiences such that for each there is an internal state of body $B_1$ causally

related to it in the proper way. If description (i) fails to refer in this way, then clearly adding to (i) the condition that $B_1$ is the body of $S_1$ will do nothing to choose between the experiences satisfying (i). But description (g) is just description (i) with this extra condition added. Hence, if (i) fails by virtue of there being more than one experience satisfying it, then (g) will fail for the same reason.

On the other hand imagine that (i) fails because there is no experience satisfying it. It follows logically from the fact that no experience satisfies (i) that no experience will satisfy (h), for both of the conditions mentioned in (i) are mentioned also in (h). For any two conditions, what does not satisfy those two conditions will not satisfy those two conditions in conjunction with a third. My conclusion then: description (i) can fail in only two ways; and for each of these ways in which (i) can fail, (g) fails also. It follows from this that if (g) succeeds in referring to an experience, then (i) succeeds in referring to an experience also.

Therefore, if (b) is true, (i) can be used to refer to an experience. But description (i) makes no reference to a person. In fact, it is just description (g) with the reference to a person dropped out. Hence, if (a), (b), and (c) are true, it is possible for (g) to identify an experience, and it is possible for (i) to identify an experience. But this is to identify an experience without reference to a person. Hence, (b) is false. Therefore, if (a), (b), and (c) are true then (b) is false.

I turn now to showing that if (a), (b), and (c) are true, then (a) is false. We have established that it is possible to identify an experience by means of description (i). In so identifying, no reference would be made to a person. Suppose that the experience so identified is a particular experience $E_1$. As Ayer says, it is necessarily true that for each experience it belongs to one and only one person. Given the necessity of this proposition, and the fact that we have identified an experience $E_1$, and that by hypothesis $S_1$ is the person to whom $E_1$ belongs,

we can identify $S_1$ by the following description: (j) the person to whom $E_1$ belongs. Given that we can identify person $S_1$ by description (j), we have identified a person without reference to his body, solely by reference to an experience of his. Therefore, if (a), (b), and (c) are true it is possible to identify a person by description (j). And if it is possible to identify a person by description (j), then (a) is false.

Ayer could answer my contention that the falsity of (a) follows from (a), (b), and (c) by saying that he meant to use the expression "identified by means of" in such a way that it is transitive. Therefore, he could claim, experience $E_1$ was identified by means of a body, and $S_1$ by $E_1$. Therefore, the identification of $S_1$ by description (j) was an identification of a person by means of a body. If Ayer were to maintain that the expression "identified by means of" is transitive, then I must give up my contention that if (a), (b), and (c) are true then (a) is false. This reply would not affect my contention that if (a), (b), and (c) are true then (b) is false. I would still have one conclusion left, at least temporarily. If Ayer makes this reply, then we once again see trouble arising because he has made no attempt to elucidate these key concepts in his position, and we do not know, therefore, whether the relation *identifies by reference to* is transitive.

3

In this section I should like to raise a question that Ayer must be able to answer if his view is to be considered tenable. In raising it, I shall be making the assumption that he accepts proposition (e). I do not know that he does but he does hold that a person can be identified by reference to his body. And it does seem reasonable to assume that what this means is that there is some relation R such that it can be known that one and only one person has R to some given body, and that the identification occurs by means of a definite description of the form "the one and only one person who has R to such and such a body." I might identify a person by means of his

automobile. And to do this I might identify him as the one and only one person who has the relation of owning to that automobile.

Assuming, then, that Ayer accepts proposition (e), let us imagine (e) instantiated for particular persons $S_1$ and $S_2$, and some particular body $B_1$. Then let us assume that the antecedent of the instantiated form of (e) is true. We will be assuming that a particular person $S_1$ does identify some particular person $S_2$ by reference to some particular body $B_1$. From this assumption together with proposition (e), it follows that $S_1$, the person doing the identifying, knows that there is one and only one person who has the relation R in question to $B_1$. The question I wish to raise is this: from what kind of evidence could $S_1$ infer that there is one and only one person who has R to $B_1$?

One thing seems clear. It could not be that $S_1$ infers that there is one and only one person having R to $B_1$ from the proposition that $S_2$ and only $S_2$ has R to $B_1$. That is, $S_1$ could not infer that there is one and only one person having R to $B_1$ from a proposition that asserts of a particular person that he has R to $B_1$. For, according to proposition (d), to know that such a proposition is true of a particular person, $S_1$ must be able to identify that particular person. But for $S_1$ to identify that particular person there would have to be another relation R′ such that $S_1$ knows that one and only one person has R′ to *some body*. Thus, an infinite regress seems to follow, given (d) and (e), from the contention that $S_1$ knows that there is one and only one person who has R to $B_1$ from evidence consisting of statements about particular persons.

We might imagine that his evidence has the following form: $B_1$ has P, and for every B: (If B has P *then* there is one and only one person having R to B). These two premises will entail that one and only one person has R to $B_1$, and hence could serve as the evidence of $S_1$, provided it is possible for $S_1$ to get evidence for them. But what kind of evidence can he

have for the generalization, other than instances of particular bodies and particular persons related by relation R? The problem comes to this. In order for $S_1$ to identify $S_2$, by reference to $B_1$, he must know of some relation R that one and only one person has R to $B_1$. The question that is puzzling is how $S_1$ could have evidence for such a proposition without ever using a proposition about a particular person.

*4*

In this last section, I should like to make a claim for which I can give no argument. Possibly, the main contention of Ayer's paper is proposition (c). Though I know of no argument to show it, proposition (c) appears to be simply false. Ayer agrees that what internal states of my body are causally necessary for my experiences is a contingent matter to be investigated by scientists. It seems equally true that whether any internal states of my body at all are causally necessary for my experiences, is a contingent matter to be investigated by scientists. For example, it seems logically possible that scientists might have discovered that in the air immediately adjacent to my body, little atomic events occur which are the immediate causally necessary conditions of my experiences; and that all of the internal states of my body are causally irrelevant to the existence of my experiences. For example, there appears to be no contradiction in the supposition that when light waves hit my eyeballs, certain events are brought about just outside of my eyeballs, which are themselves immediate causally necessary conditions of my having the experience of seeing. However, I have no argument to present. I can only put my intuition up against that of Prof. Ayer's.

Ayer does admit that an overt bodily movement of someone other than myself, e.g., the waggling of a finger by the man standing next to me, might be an immediate causally necessary condition of some thought of my own. For this reason, he puts the word "internal" into proposition (c). It seems absolutely astonishing that a philosopher would think

it logically possible for someone's waggling of his finger to be an immediate causally necessary condition of my having some thought, but at the same time think it logically impossible for some internal state, e.g., his flexing certain muscles or his heart skipping a beat, to be so related to a thought of mine. It seems clear that whatever is true in the one case is true, also, in the other; and that in both cases it is a contingent fact that such bodily states of other persons are or are not immediate causally necessary conditions of my thoughts and experiences.

I suppose that a philosopher can be justified on holding certain propositions true even though they seem paradoxical or even false, provided no argument is known against their truth. But when confronted with propositions that seem so plainly false, it would be comforting to have at least *some* arguments in their favor.

# REJOINDER

A. J. Ayer

I SHALL BEGIN by trying to meet Mr. Gettier's request for a clearer explanation of what I mean by saying that one type of thing is identified by reference to another. In the case of anything of which it can significantly be said that it comes into existence at a particular time and remains in existence for a time, however short, I assume that there must be some way in which its presence can be detected. This is not to say, of course, that the presence of any individual object is detectable by any observer at any time. Clearly I cannot now detect the presence of objects which have ceased to exist or not yet come into existence; and among the objects which do now exist there are many whose presence I have no means of detecting because they are too remote in space for me to have access to them. But if the reference that I make to any such object is to be intelligible there must at least be a standard method of detecting the presence of objects of that kind. There are also, I think two other conditions that have to be satisfied. There must be some criterion by which one member of a given class of objects is differentiated from another; and in the case of things which have more than a momentary duration there must be some criterion of self-identity. There must be some way of determining whether two episodes which are separated in time do or do not belong to the history of the same object.

Now though everything which exists in time must be ac-

cessible to observation in some sense or other, the mode of access need not always be straightforward. There are types of objects, or occurrences, the presence of which is detectable only through detecting the presence of objects or occurrences of different types. Thus, one of the standard means of detecting the presence of electrons is to observe characteristic marks on photographic plates; unconscious states of mind are typically manifested by certain forms of behavior; the actions of corporate bodies, like parliaments or armies, are revealed only in the actions of the men and women who compose them. Of course I am not suggesting that the relation between the different types of entities is the same in each of these examples: obviously the relation of a particular soldier to the army of which he is a member is not of the same order as the relation between a particular set of marks on a photographic plate and the electrons of which they are the traces. But what these examples do have in common is that in all of them there is a distinction between something which is manifested and something of a different type through which this manifestation is effected. And this is what I had in mind when I spoke of one thing's being identified by reference to another.

In the case of persons, their bodies, and their experiences, there is, however, a complication which I hoped I had brought out in my essay. Gettier attributes to me the view that experiences can be identified only through the identification of the persons whose experiences they are, and that persons can be identified only by reference to their bodies. Though I believe this to be true of the experiences which one ascribes to other persons, I do not believe this to be true of one's own experiences, as I hoped that my essay had made clear. Consequently, while I hold that the identification of the person whose experience it is can in the appropriate circumstances be a sufficient condition for identifying an experience, I do not hold that it is necessary. This does not mean that I think there could be experiences which were not the

experiences of any person. It is just that in one's own case I admit the possibility of one's having what might be called direct access to them.

On the other hand, I do hold that a person can be identified only by reference to his body, even in the special case in which the person is oneself. My reason for this is that even in referring to oneself as a person one has to rely upon some criterion of self-identity; and I do not see how personal identity is to be analyzed except in terms of the identity of a body. I admit that we do make use of other criteria but argue that they are parasitical upon the criterion of bodily identity; I do not think that they could sustain a concept of personal identity on their own. Thus I reject the idea of there being disembodied persons, even as a logical possibility, because I do not see how the experiences of such a person could be held together; that is, I do not see what relation there could be between them in virtue of which they would be the experiences of one and the same person.

In the light of these explanations, I will make some brief comments on the several points which Gettier raises. Of the five propositions which he begins by listing, I accept (a) (b) and (e) and a weaker version of (d), in the sense that I hold that if S is intelligently to assert or judge that fx, x must at least be the sort of thing that he knows how to identify. I do not, however, accept (b) if this is taken to imply that experiences are necessarily identified by reference to the persons who have them.

It should be clear from what I have already said that the first of Gettier's apparent counterexamples is not in fact a counterexample to proposition (a) as I interpret it. There is, indeed, a sense of the word "identify" in which I can properly be said to have identified something if I have given a description which applies to it and to nothing else: and in this sense I might be said to have identified A by reference to B if the description which I had given of A was one in which B was

mentioned as standing to A in some unique relation. Thus, in Gettier's example, Einstein is identified by reference to the theory of relativity, as being the person who first thought of it. Hence, if this had been the sort of identification that I had in mind, Gettier would have found a counterexample to my thesis and could easily have produced many others. But I hope that I have now made it clear that the sort of identification which I did have in mind was not one to which such examples would be relevant.

The second of Gettier's counterexamples is a counterexample to proposition (b). But since I do not hold proposition (b) this does not affect me.

In the second part of his paper Gettier develops an elaborate argument to show that, if my theory is correct, an experience E can be uniquely characterized as an experience of a given type which has an internal state of a body $B_1$ for its immediate necessary condition, and he infers from this, first that on my own principles experiences need not be identified by reference to persons since they can be identified by reference to bodies, and secondly, that persons need not be identified by reference to their bodies since they can be identified by reference to their experiences. I have no fault to find with the argument except that to make it entirely rigorous he should have included a reference to the time at which E occurs, but the inferences which he draws from it do not disturb me. I do not claim immunity on the ground that I reject proposition (b) since I should wish to maintain it with respect to the experiences of persons other than oneself, which, if Gettier were right, I could not consistently do. The point is rather that for me the distinction which he has to make between identifying an experience by reference to a person and identifying it by reference to that person's body presents a false antithesis. In my view, as I thought my essay would have shown, to identify an experience as the experience of a particular person is equivalent to identifying it by reference

to the body by which the person in question is himself identified. If I speak of identifying experiences by reference to persons rather than by reference to bodies, it is only because we commonly talk of persons rather than of bodies as having experiences. But I never suggested that to identify an experience by reference to a person was a distinct process from identifying it by reference to a body. And for the same reason I do not count the fact that a person may be uniquely characterized as the owner of an experience which is itself identified by reference to a body as at all an objection to my thesis that persons are identifiable only by reference to their bodies. In this case, therefore, the question whether my relation of "identifies by reference to" is transitive does not arise, but in fact I do hold it to be transitive.

The same mistake seems to me to vitiate the argument which Gettier seeks to develop in the third part of his paper. He maintains that I am committed to holding that if a person $S_1$ identifies some other person $S_2$ by reference to a particular body $B_1$, $S_1$ must know that there is one and only one person who stands to $B_1$ in the relation R on which the identification is based; and he then argues that "it could not be that $S_1$ infers that there is one and only one person having R to $B_1$ from the proposition that $S_2$ and only $S_2$ has R to $B_1$." If I follow him correctly, his ground for this is that $S_1$ could not attach sense to the proposition that $S_2$ has R to B unless he had independently identified $S_2$, and that, on my principles, the only way in which $S_1$ could identify $S_2$ independently of knowing that $S_2$ bore the relation R to a particular body B, would be by knowing that $S_2$ bore the relation R to some body or other; and he understandably fails to see how one could come to know that a person bore the relation R to some body or other except through coming to know that he bore this relation to a particular body. But the sufficient answer to all this is that, given that the body in question has the characteristics of a person, as opposed, say to an animal or an inanimate

thing, the identification of the person just consists in the identification of the body. There is no question of the person's having to be independently identified. Of course if the statement that the body is the body of the person whose body it is, is not to be merely trivial, the person will have to be described in some other way than merely as the owner of that body. But, as I have already explained, the fact that a person may be uniquely described in ways that make no explicit mention of his body is not incompatible with the thesis that it is only by reference to their bodies that persons are, in my sense, identifiable.

I may remark in passing that while I do make it necessary that one and only one person "inhabits" a given body at a given time, this rule could easily be relaxed if it were found advisable. At present I do not allow for any other means of describing cases of alleged co-consciousness than by saying that they are cases of one person with a split personality. But it would not be hard to introduce subsidiary criteria of personal identity which would make it significant to talk of there being two or more persons simultaneously inhabiting a single body. However, as in the other para-normal cases with which I dealt, these criteria would still be parasitical upon the primary criterion of physical identity.

Finally, Gettier states that he just does not find it credible that the relation in virtue of which a person's experiences are ascribable to him should consist in their dependence upon some internal state of his body as their immediately necessary causal condition; and he complains that if I was going to advance a proposition which seems so plainly false, I ought at least to have produced some arguments in its favor. My answer to this second charge is that I did produce a number of arguments which seemed to me to leave no alternative to the view that a person's ownership of his experiences consists in their standing in some causal relation to his body; and the theory that this relation was that of their having some internal

state of his body as their immediately necessary causal condition seemed to me to fit the facts better than any other that I could think of. I do not deny, however, either that it seems excessively artificial or that it forces me to attribute necessity to propositions which I might otherwise have been more inclined to regard as contingent. It may well be then that on this point Gettier's intuition is to be trusted. My own intuition goes no further than the belief that even if this theory is not the answer, some theory of a similar type is most likely to be true.

## NOTES ON AYER ESSAY

1. P. F. Strawson, *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen & Co. Ltd., 1959).
2. *Ibid.*, p. 102
3. *Ibid.*, p. 99.
4. Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Basil Blackwell, 1953), p. 293.
5. G. E. Moore, *Philosophical Studies* (London: Kegan Paul, Trench, Trubner and Co., Ltd., 1922), pp. 31–96.
6. *Op. cit.*, p. 108.
7. *Ibid.*, p. 111.
8. Stuart Hampshire, *Thought and Action* (London: Chatto and Windus, 1959), pp. 47–8.
9. *Ibid.*, p. 75.
10. *Op. cit.*, p. 110.
11. C. K. Ogden and I. A. Richards, *The Meaning of Meaning* (London: Kegan Paul, Trench, Trubner & Co., Ltd.; New York: Harcourt, Brace and Co., Inc., 1938).
12. A. J. Ayer, *The Problem of Knowledge* (London: Macmillan and Co. Ltd.; New York: St. Martin's Press, 1958), p. 222.
13. *Loc. cit.*
14. *Ibid.*, p. 226.