

Strawson's Argument against the Cartesian Thesis:  
An Interpretation

Robert L. Latta

In Chapter 3 of his book *Individuals*, the much-discussed chapter "Persons," P. F. Strawson puts forward an argument which he summarizes in these words:

... One can ascribe states of consciousness to oneself only if one can ascribe them to others. One can ascribe them to others only if one can identify other subjects of experience. And one cannot identify others if one can identify them *only* as subjects of experience, possessors of states of consciousness.<sup>1</sup>

This is one of the central arguments in "Persons." I shall refer to it as Strawson's argument against the Cartesian thesis—with apologies to Descartes if I misapply his name. By 'the Cartesian thesis,' I mean the thesis that the experiences of a human being pertain to an entity to which nothing but experiences pertain, and his or her physical attributes to a numerically distinct, i.e., a second, entity. What I call the Cartesian thesis does not, of course, constitute a full Cartesian theory of mind and body, for that would include, among other things, an account of the relations between a human being's mind, a subject of experience, and his or her body, a subject of physical attributes. The Cartesian thesis does, however, constitute the core of any Cartesian theory of mind and body, and it is this core thesis that Strawson attacks.

Though it is not typical of his writing, in *Individuals*, Strawson leaves himself open to indefinitely many widely diverging interpretations. At certain critical points, all the reader has to go on are short remarks that resemble nothing so much as notes of the sort that one makes for one's own use alone. Not surprisingly, Strawson's critics have interpreted the argument I have quoted in various ways. Many believe themselves to have refuted that argument. What they have in fact refuted, however, is a remarkably diverse group of arguments as numerous as the critics themselves, arguments which for the most part, if the truth be told, are quite easy to refute. My thesis is this: Strawson's argument against the Cartesian thesis can be given an interpretation faithful to the text under which it is formidable and, if sound, philosophically significant, and under which the great majority if not all of the criticisms that appear in the published literature simply

## Strawson's Argument against the Cartesian Thesis: An Interpretation

do not apply. I shall defend this thesis by presenting an interpretation of my own. I shall appeal to Strawson's wording to show that my interpretation is accurate with respect to those points on which he is explicit. I shall cite a few of his critics, but only incidentally, in order to explain my own readings of certain passages more clearly.

The paper is divided into several sections. The first contains remarks on the meanings of certain key terms. In the second, I present and defend my interpretation of Strawson's argument against the Cartesian thesis, and in the third, I discuss the implications of that argument for that thesis. In the fourth section, I take up Strawson's thesis that "the concept of a person" is "logically primitive," for his argument against the Cartesian thesis constitutes his defense of the more problematical clause of that thesis. In the fifth section, I present and reply to several objections that have occurred to me at various times. The sixth and final section contains a brief summary.

### 1. Remarks on Terminology

Below, I make remarks on certain key terms. These remarks are meant to clarify, but they are not in every case offered as full definitions.

(1) When Strawson uses the term 'ascribe' with 'experiences,' or 'a pain,' or 'feelings' or the like as direct object, he means 'to assert to belong or pertain,' and not 'to assign to a cause.' In this application, I shall use this term in the same sense.

(2) Strawson uses the expression 'a particular' to mean a particular thing, event, state, place, or time, etc., where particulars stand in contrast with various sorts of non-particulars, like universals, species, numbers, attributes, etc.<sup>2</sup> I shall adopt the expression 'a particular' and shall use it in this way. It must be admitted that this expression is in a way problematical: it is problematical in that it is not possible to give an accurate and full explication of its meaning without first solving the large and fundamental problem of the particular/general dichotomy. But this much at least is, I think, clear: leaving out of account the problematical case of particular times and places, if an individual, *X*, is to be considered a particular, then it must be possible in principle that there are individuals numerically distinct from *X* but exactly like it. This condition is met, for example, by the pen which I am holding, and by tokens of the numeral '3,' but not by sobriety, or by the number three.

(3) Strawson applies the expression 'identification of particulars' to several different operations. He applies it, first, to the process of establishing references to particulars in speech, i.e., in interpersonal discourse. In this first application, if a speaker attempts to establish a reference to a particular and his hearer understands that reference, i.e., knows

what particular the speaker is referring to, then the hearer, and hence the speaker, are said to have identified the particular in question. A hearer, then, identifies a particular in the sense that he understands a speaker's reference to it, whereas a speaker identifies one in the sense that he makes such a reference to it that his hearer understands that reference.<sup>3</sup> At a later point, Strawson applies this expression to a somewhat different operation: the process of establishing references to particulars "in thought," i.e., in the sort of intrapersonal discourse that is characteristic of solitary thought. I shall use the expression 'identification of particulars' myself in the senses defined by Strawson, although the expression 'establishment of references to particulars' would, I believe, do just as well.

There is an alternative interpretation of Strawson's use of the term 'identify' that I wish to mention. It appears that according to Donald Sievert, in Strawson's usage, 'identify' means 'locate in space and time.'<sup>4</sup> This, I think, is a serious misunderstanding. In explaining what meaning he gives to the term 'identify,' Strawson makes no reference to space and time. Reference to spatial-temporal location occurs only in his treatment of the question of the ways in which references to particulars can be established. Defining the term 'identify,' and taking up this question, are two entirely distinct things. One must do the former before one can do the latter.

(4) By 'a subject of experience,' or for short 'a subject,' I mean simply an entity to which experiences can with truth be ascribed. The question in what sense or way experiences are ascribable to the entities which are said to have them has a central place in Strawson's case against empiricist accounts of the idea of a subject of experience, but is not closely connected with his argument against the Cartesian thesis. I shall therefore avoid this question in this paper in order not to obscure the issues at hand.

(5) By 'a subject of physical attributes,' I take Strawson to mean, and shall mean myself, simply an entity to which physical attributes can with truth be ascribed. Thus, a stone, a tree, and the dead body of Jones all count as subjects of physical attributes. Since the term 'subject' is not often used in this connection, confusion can arise. For example, R. L. Phillips states that Cartesian dualism, as described by Strawson,

... regards a person as a compound of 'two subjects of experience', a mind and a body....<sup>5</sup>

Strawson means, however, not two subjects of experience, but a subject of experience and a subject of physical attributes.

(6) By 'a pure subject of experience,' I mean an entity to which experiences and

nothing but experiences can be ascribed. Strawson sometimes uses the term 'a pure individual consciousness' to refer to entities of this type.

(7) In order to explain more clearly what Strawson means by 'the concept of a person,' I shall first explain what he means by 'a logical type.' He does not define this latter term, but the following definition of it appears accurately to reflect his usage: a logical type is a type defined as the type to which predicate-expressions of a certain class or classes are significantly applicable. Not all types of particulars are logical types. Consider this definition of the term 'pulley':

... a sheave with the pin on which it turns, the frame in which it runs, and the flexible rope, cord, or chain passing through the groove that is used singly to change the direction and point of application of a pulling force applied at one end of the rope, cord, or chain and singly or in any of various definite combinations to increase the applied force esp. for lifting weights...<sup>6</sup>

To define the term 'pulley' thus is not to define the type of particulars in question in the way just described. Pulleys constitute a type, but, at least by the definition above, not a logical type.

In order to understand Strawson's argumentation, it is essential to realize that he uses the term 'the concept of a person' in a special, technical way:

... What I mean by the concept of a person is the concept of a type of entity such that *both* predicates ascribing states of consciousness *and* predicates ascribing corporeal characteristics, a physical situation etc. are equally applicable to a single individual of that single type...<sup>7</sup>

By this definition, persons constitute a logical type. Rocks, for example, are not of this particular logical type, since experiences are not ascribable to rocks. Pure subjects of experience are not of this type because physical attributes are not ascribable to them. But *if* the reasoning that I shall attribute to Strawson is sound, then Strawson himself while he lives and breathes, Descartes while he lived and breathed, and my neighbor's pet canary while it lives and breathes, count as persons by Strawson's definition, for if the reasoning in question is sound, then as long as these three entities function biologically in normal ways, they are to be conceived as subjects of both experiences and physical attributes.

It should be noted that Strawson's definition of the concept of a person is a stipulation, not a lexical definition. Thus, although, evidently, he intends that his definition accord with common usage to the extent that if his reasoning concerning the concept of a subject

of experience is sound, then any person in the most usual sense counts as a person by it, he does not put it forward as a description or analysis of common usage. His use of the term 'a person' has brought about confusion; he might have done better to use a term like 'a subject of both experiences and physical attributes' instead, for this is exactly what he means by 'a person.' In this paper, I shall use the term 'a person' in the sense Strawson defines, but with caution.

(8) By 'a disembodied person,' Strawson means a subject of experience that once had physical attributes but has lost them. To put it another way, a disembodied person is a pure subject of experience which in an earlier stage of its existence was a "person" in the technical sense, but then lost its physical aspect. He also uses the term "former person" to refer to such an entity.

(9) By 'a Cartesian ego,' I mean an entity which is a pure subject of experience throughout the duration of its existence. The types 'Cartesian ego' and 'disembodied person,' then, are mutually exclusive, though both a Cartesian ego and a disembodied person are to be understood to be pure subjects of experience.

## 2. Strawson's Argument

The first premiss of Strawson's argument against the Cartesian thesis reads as follows:

... It is a necessary condition of one's ascribing states of consciousness, experiences, to oneself, in the way one does, that one should also ascribe them, or be prepared to ascribe them, to others who are not oneself...<sup>8</sup>

When he comes to summarize his argument, he states this premiss more simply:

... One can ascribe states of consciousness to oneself only if one can ascribe them to others.<sup>9</sup>

Strawson comments on and defends this premiss of his argument in a very compact and difficult three-paragraph footnote. In the first paragraph, he brings an objection:

... I can imagine an objection to the unqualified form of this statement, an objection which might be put as follows. Surely the idea of a uniquely applicable predicate, i.e. a predicate which belongs to only one individual, is not absurd. And, if it is not, then surely the most that can be claimed is that a necessary condition of one's ascribing predicates of a certain class to one individual, i.e. oneself, is that one should be prepared, or ready, on appropriate occasions, to ascribe them to other individuals, and hence that one should have a conception of what those appropriate occasions for ascribing them would be; but not, necessarily, that one should actually do so on any occasion.<sup>10</sup>

Clearly, the "unqualified form" of the premiss under consideration is the form obtained

by deleting the qualifying phrase 'or be prepared to ascribe them': 'It is a necessary condition of one's ascribing experiences to oneself, in the way one does, that one should also ascribe them to others.' The objection that Strawson brings against the unqualified form of the first premiss of his argument, then, is this. The idea of a predicate-expression that applies truly to only one individual is perfectly intelligible. Hence, it is conceivable that those predicate-expressions that are used to ascribe experiences should be truly applicable to only one individual. But clearly, it is conceivable that that individual, who would of course be the only subject of experience in the universe, should apply those predicate-expressions to himself but not to any other individual. He might often say, for example, "I feel lonely," but never "It feels lonely." Hence, it cannot be said that a subject of experience can ascribe experiences to himself only if on occasion he actually ascribes experiences to others. Therefore the premiss under consideration, in its unqualified form, is false.

Strawson allows this objection to stand. He maintains only the weak version of the first premiss of his argument, the version according to which, to be in a position to ascribe experiences to oneself, one need not actually ascribe them to others, but need only "be prepared" to do so "on appropriate occasions" (see the last quotation above). As I read him, his phrase 'be prepared to ascribe experiences to others on appropriate occasions' means 'know under what conditions it is correct to apply the relevant predicate expressions to others.' His defense of this assertion is to be found in the third paragraph of his footnote:

The main point here is a purely logical one: the idea of a predicate is correlative with that of a *range* of distinguishable individuals of which the predicate can be significantly, though not necessarily truly, affirmed.<sup>11</sup>

By "main point," he means 'underlying point.' He takes the position that the proposition in question entails the first premiss of his argument. This proposition appears to be nothing more than the commonplace that an expression like '... is angry,' or '... is tall,' or '... is made of rosewood' can be applied to any number of individuals. To put this more accurately: it is an essential characteristic of predicate-expressions that, in any one of its meanings, any given predicate-expression can in principle be applied significantly to any number of individuals of certain types. Thus, to illustrate, the meaning of the predicate-expression '... is angry' is such that in principle, it can be applied significantly to any number of people, chimpanzees, dogs, etc. Moreover, in a different sense, doubtless a derivative one, it can in principle be applied significantly to any number of rivers or

storms.

Strawson leaves it to his reader to work out the connection between his point about the logic of predicate-expressions and the first premiss of his argument against the Cartesian thesis. Note that the former is more general than the latter. His point about predicate-expressions has to do with predicate-expressions in general, whereas the first premiss of his argument has to do only with predicate-expressions of a certain class: those that are used to ascribe experiences. I shall now assume the former and derive the latter from it. The first explicit step in the derivation is to assume the negation of the consequent of the conditional to be derived: Assume that one cannot ascribe experiences to others in that one does not know under what conditions it is correct to apply the relevant predicate-expressions to others. Then from one's own point of view, as a matter of principle, expressions that are used to ascribe experiences can be applied significantly to only one individual—viz., oneself. But it is an essential characteristic of predicate-expressions that any given one can in principle be applied significantly to any number of individuals. Consequently, in the case in question, expressions that are used to ascribe experiences cannot function *as* predicate-expressions. But this is to say that they cannot be used to ascribe experiences, which in turn is to say that there are no expressions that can be used for this purpose. But if this is so, then, obviously, one cannot ascribe experiences to oneself, in that one does not have the requisite predicate-expressions at one's command for use in speech and thought. Therefore, to convert to the contrapositive, if one can ascribe experiences to oneself in that one has the requisite predicate-expressions at one's command for use in speech and thought, then one can ascribe experiences to others also, in that one knows under what conditions it is correct to apply those predicate-expressions to others.

This derived proposition represents my interpretation of the first premiss of Strawson's argument against the Cartesian thesis. It is, if true, an inobvious truth that carries considerable philosophical significance. I shall now make a short digression to describe one way in which it is significant. Many philosophers have pointed out a peculiar fact about predicate-expressions like, for example: '... (suffer / suffers) a twinge of regret,' '... (feel / feels) dizzy,' etc. In one special case, the case in which one affirms 'I am suffering a twinge of regret' or 'I feel dizzy,' etc., one can affirm of a person, viz., oneself, that he or she is having an experience of such and such a sort without having to appeal to behavioral criteria, and without having to ask another person, to determine whether one's affirmation is justified. To put this more concisely, the criteria for the

ascription of experiences play no part in the process of self-ascription of experiences. This fact might quite naturally give rise to the following impression: it is conceivable that a subject should ascribe experiences to himself, and yet have no conception of *criteria* for ascribing them, i.e., to others. But if the proposition which figures as the first premiss of Strawson's argument against the Cartesian thesis is true, then this impression is mistaken, for if that proposition is true, then, if a subject had no conception of criteria for ascribing experiences to others, he would have at his command no linguistic expressions to use to ascribe experiences to any subject, even himself.

In the critical literature, the first premiss of Strawson's argument is given many different interpretations. There is one in particular that I wish to mention, for it is very popular, and I want to distinguish it sharply from my own. According to this interpretation, the premiss in question can be put as follows:

One can ascribe experiences to an entity that one conceives of as oneself only if one can ascribe experiences to entities that one conceives of as others.<sup>12</sup>

Under my interpretation, the premiss in question is fundamentally different from this, and can be put this way:

One can ascribe experiences to a certain entity—viz., the one which in fact one conceives of as oneself, only if one can ascribe experiences to other entities in that one knows under what conditions it is correct to ascribe an experience of this or that sort to another.

The latter proposition, given that the condition stated in its consequent is not met, rules out even the possibility that one should ascribe experiences to a certain entity, viz., oneself, while failing to conceive of that entity as oneself. But the former proposition, given that the condition stated in its consequent is not met, does not rule out this possibility. In this way, then, the latter proposition expresses a more sweeping thesis than the former. But the consequence to be drawn from Strawson's point about the logic of predicate-expressions is this latter, more sweeping thesis.

The second premiss of Strawson's argument reads as follows:

... One can ascribe [states of consciousness] to others only if one can identify other subjects of experience.<sup>13</sup>

This might be taken to mean: as a practical matter, if one cannot establish references to others, then one cannot ascribe experiences to them. Quite clearly, this premiss so interpreted is true, for so interpreted, it follows from this general principle: ascribing a characteristic or a state to a particular essentially involves referring to that particular.



Thus, for example, ascribing the characteristic of being red to the barn involves referring to the barn—by means of an expression like ‘the barn,’ for example, or perhaps by means of a demonstrative like ‘that’ coupled with a gesture.

I think, however, that the thesis Strawson has in mind is a somewhat more subtle one. The second premiss of his argument should of course dovetail with the first and third premisses. The first premiss, as I have explained it, is this: If *A* (a proposition there is no to repeat here), then *B*: one can ascribe experiences to others in that one knows under what conditions it is correct to apply the relevant predicate-expressions to others. I shall take the antecedent of the second premiss to be this proposition *B*. But the third premiss, as I interpret it, is this: If *D* (a proposition I shall introduce later), then not-*C*: one cannot identify others at all, in that no coherent account of how to go about identifying them can possibly be formulated. I shall take the consequent of the second premiss to be proposition *C*, since the negation of the consequent of the second premiss is the consequent of the third premiss: see Strawson’s summary of his argument, quoted on page 1 above. As I interpret it, then, the second premiss of Strawson’s argument is this: if one can ascribe experiences to others in that one knows under what conditions it is correct to apply the relevant predicate-expressions to them, then one can identify others in that a coherent account of how to go about identifying them can be formulated.

As far as I can see, Strawson does not indicate how this premiss might be defended. The case for it, however, can, I think, be stated rather simply: To say that no coherent account of how to go about identifying subjects of experience numerically distinct from oneself can possibly be formulated is to say that no question at all arises of identifying others. But if no question at all arises of identifying others, then there is no such thing as knowing under what conditions it is correct to ascribe an experience of this or that sort to another subject. From this it follows that if one does know under what conditions it is correct to ascribe an experience of this or that sort to another subject, then it is possible to formulate a coherent account of how to go about identifying others.

The third premiss of Strawson’s argument against the Cartesian thesis is this:

... One cannot identify others if one can identify them *only* as subjects of experience, possessors of states of consciousness.<sup>14</sup>

I take his meaning to be this: if, in identifying subjects of experience numerically distinct from oneself, all one has to go on, of all that is ascribable to them, are their experiences, then one cannot identify others at all, in that no coherent account of how to go about identifying them can possibly be formulated. Or, to put the point more shortly: if

subjects of experience are Cartesian egos, then one cannot identify subjects numerically distinct from oneself. It is possible to piece together an argument for this premiss from things Strawson says in various passages. I shall do this now.

To begin with, Strawson asserts that any particular that has a spatio-temporal position has a unique spatio-temporal position, i.e., one shared by no other particular.<sup>15</sup> But clearly, if each Cartesian ego has a unique spatio-temporal position, then, at least in theory, there is no problem about how to identify them. Clearly, then, it is Strawson's view that Cartesian egos are not to be conceived as having spatio-temporal positions. Human bodies do, of course, have spatio-temporal positions, but Strawson maintains further that there is no guarantee of a logical or conceptual sort that given a human body, only one Cartesian ego is causally associated with it. In order to make this point, he puts a rhetorical question:

...Can we not identify [a Cartesian ego] as... 'the subject of those experiences which stand in the same unique causal relation to body N as *my* experiences stand in to body M'?...<sup>16</sup>

He makes two replies to this question, one of which reads as follows:

...What right have we, in this explanation, to speak of *the* subject, implying uniqueness? Why should there not be any number of subjects of experience—perhaps qualitatively indistinguishable—each subject and each set of experiences standing in the same unique relation to body N (*or* to body M)? Uniqueness of the body does not guarantee uniqueness of the Cartesian soul.<sup>17</sup>

His meaning, I take it, is the following. Assume that subjects of experience are Cartesian egos. Then it is possible in principle that each of many numerically distinct sets of experiences, perhaps of exactly the same description, each had by a numerically distinct subject, is causally connected in certain ways with the physical states of a given human body *B*. That is to say, this is a possibility in that nothing about the concept of a Cartesian ego rules it out as a possibility. But it follows that there can be no warrant for the employment of definite descriptions of the forms 'the set of experiences that is causally connected in such and such ways with the physical states of body *B*' or 'the Cartesian ego that has the set of experiences that is causally connected in such and such ways with the physical states of body *B*.'

Furthermore, in Chapter 1 of *Individuals*, Strawson puts forward a complex account of the ground of the possibility of identifying particulars. As I read him, his main conclusion is the following. People can in fact identify particulars of a vast range and

of many different types in speech and in thought, and can in principle identify particulars of a yet more extensive range and further types. Of these particulars, some but not all have spatio-temporal positions. Particulars that have spatio-temporal positions form a system of relations within which each such particular has a unique position. This is the only system of relations that can serve as the core of a system within which each particular, of the very extensive range and many types which people can in principle identify, has a unique position. But this wider system of relations, the one within which each particular of whatever type has a unique position, is an indispensable instrument in the identification of particulars. For the only way to distinguish a given particular from absolutely all others is to indicate or allude to some circumstance in virtue of which it has a unique position within this system. Thus, to illustrate, if a certain piece of land is the only one owned by a certain man, then it has a unique position within the system of relations in question in that it has such and such a position in space and time, and in addition, in what is perhaps a figurative sense of 'position,' it has a unique position within this system in that it is the only piece of land owned by that man, and *he* has a unique position, or series of positions, within the system.

But taken together, these views suggest the following argument. Assume that subjects of experience are Cartesian egos. Then they have no positions in space. Moreover, given any particular in space, there is no relation such that there is ground for assurance in the logic of the matter that that relation obtains between that particular on the one hand, and only one subject of a given general description on the other. Hence, subjects of experience do not have positions within the general system of relations among particulars. Hence, one cannot have any conception of how to fit particulars of this type, subjects of experience, into this system. Hence, one has no conception of how to go about identifying subjects of experience. Therefore, if subjects of experience are Cartesian egos, then one cannot identify others, in that one cannot formulate any sound conception of how to go about doing this.—But this is the third premiss of Strawson's argument against the Cartesian thesis.

Now let us return briefly to Strawson's rhetorical question whether one can identify a Cartesian ego numerically distinct from oneself as "the subject of those experiences which stand in the same unique causal relation to body N as *my* experiences stand in to body M."<sup>18</sup> As I wrote, he makes two replies to this question. I have incorporated the second of these replies into the argument stated in the preceding paragraph, but the first reply ought to be considered too, if only because it has been widely misinterpreted.

It reads as follows:

... This suggestion is useless. It requires me to have noted ... that *my* experiences stand in a special relation to body M; but it requires me to have noted this as a condition of being able to identify other subjects of experiences, i.e. as a condition of my having the idea of myself as a subject of experience, i.e. as a condition of thinking of any experiences as *mine*...<sup>19</sup>

Few of Strawson's critics recognize that in this passage, he appeals to the first and second premisses of his argument against the Cartesian thesis to show that the suggestion in question, in relying upon the expression 'my experiences,' presupposes what it purports to establish—viz., the possibility of identifying Cartesian egos numerically distinct from oneself. His reasoning can be restated as follows. If one is in a position to identify a Cartesian ego numerically distinct from oneself by means of the expression 'the subject of those experiences which stand in the same unique causal relation to body N as my experiences stand in to body M,' then, of course, one is in a position to ascribe experiences to oneself. So, in relying upon this expression, the suggestion in question presupposes that one can ascribe experiences to oneself. But by the first premiss of the argument against the Cartesian thesis, if one can do this, then one can ascribe experiences to Cartesian egos numerically distinct from oneself. But in turn, by the second premiss of that argument, if one can do this, then one can identify Cartesian egos numerically distinct from oneself. So, in presupposing that one can ascribe experiences to oneself, the suggestion in question presupposes that one can identify Cartesian egos numerically distinct from oneself. But this, of course, is just what it purports to establish.

As I interpret Strawson's argument against the Cartesian thesis, then, its three premisses have the form: 'if A then B,' 'if B then C,' 'if D then not-C.' From this, it is possible to derive: 'if D then (not C and (not-B and not-A)).' Substitution yields the following proposition, which I shall take to be the conclusion of Strawson's argument: If, in identifying subjects of experience numerically distinct from oneself, all one has to go on, of all that is ascribable to them, are their experiences, then it is impossible in principle for one to identify other subjects, and impossible in principle for one to ascribe experiences to other subjects or to oneself.

### 3. Implications for the Cartesian Thesis

As I pointed out in the introduction, the core of any Cartesian theory of the relation between mind and body is the Cartesian thesis: the thesis that the experiences of a human being pertain to an entity to which nothing but experiences pertain, and his or her

physical attributes to a numerically distinct, i.e., a second, entity. Now if the argument that I have presented and attributed to Strawson is sound, it establishes that the Cartesian thesis is untenable. But since this might be considered less than obvious, I shall now show that it is so.

Suppose that people actually do ascribe states, characteristics, properties, etc. to particular things of a certain type *T*. Then, clearly, there exist linguistic conventions of two sorts: (1) conventions for the use of expressions to *refer* to particulars of type *T*, and (2) conventions for the use of expressions to *ascribe* states etc. that can be manifested by particulars of type *T*. Let us adopt the following definition. Let concept *C* be the concept of type *T* of particulars. Thus, concept *C* might be the concept of a certain type of automobiles. Then concept *C* is *incorporated in language* if there exist conventions for the use of expressions to refer to particulars of type *T*, and conventions for the use of expressions to ascribe states etc. that can be manifested by things of type *T*. To say that concept *C* cannot be incorporated in language, then, is to say that for one reason or another, it is not possible to institute conventions of one or the other of these two sorts.

Now let us assume for a moment that Strawson's argument against the Cartesian thesis is sound, and hence its conclusion true. Although the expression 'is incorporated in language' does not appear in Strawson's presentation, the conclusion of his argument can be put as follows: 'If we assume that subjects of experience are Cartesian egos, then the question how to introduce into discourse predicate-expressions that can be used to ascribe experiences to oneself, and in addition, the question how to institute conventions for establishing references to others, represent problems that are in principle unsolvable. But this is to say that as a matter of principle, the concept of a Cartesian ego cannot be incorporated in language.' Now Cartesian theorists, of course, present the concept of a Cartesian ego as the concept of a type of particulars that actually exist, or that can at least be conceived to exist. But to say that as a matter of principle, the concept of a Cartesian ego cannot be incorporated in language is to say that it is impossible in principle to formulate any sound conception of how to go about making statements, even false ones, about individual Cartesian egos. But if it is impossible in principle to say things about individual Cartesian egos, then, to employ a Strawsonian phrase which expresses one of Strawson's thoughts, particulars of this type cannot figure in our speech and thought about the world, and from this it follows that they cannot be conceived to exist. Therefore, if Strawson's argument is sound, the concept of a Cartesian ego is a

pseudoconcept in that particulars of the type it represents cannot even be conceived to exist.

Strawson claims that the Cartesian thesis is "incoherent,"<sup>20</sup> by which he means that it is inconsistent with at least one of its own presuppositions. I shall now explain why he makes this claim, and in doing so, I shall show once again that if his argument against the Cartesian thesis is sound, then that thesis is indeed untenable. Assume that it is impossible to ascribe experiences to oneself in that, for whatever reason, as a matter of principle, no predicate-expressions to serve the purpose can be introduced into discourse. Assume further that for whatever reason, it is impossible in principle to identify other subjects of experience. Then, of course, no question arises of ascribing experiences to any subject, either to oneself or to others. But then, of course, the question whether the Cartesian thesis is true, that is, the question whether experiences are in the final analysis to be ascribed to Cartesian egos or to entities of some other description, simply does not arise, which is to say that it cannot legitimately be raised. In this sense, the Cartesian thesis carries a presupposition: that it is possible in principle to ascribe experiences to oneself and to identify other subjects. But if Strawson's argument against the Cartesian thesis is sound, then, under the assumption that the Cartesian thesis is true, that is, under the assumption that subjects of experience are Cartesian egos, these things are not possible. Therefore, if Strawson's argument is sound, the Cartesian thesis is inconsistent with one of its presuppositions, and hence is untenable.

In his article "Cartesian Dualism and the Unity of a Mind," Chin-Tai Kim offers an interpretation that is similar to mine on the point under discussion:

...The incoherence of Cartesian dualism, according to Strawson, ...consists in the fact that whereas this theory is supposed to provide, among other things, an explanation of how self-attribution of experiences is possible, its basic thesis that a subject of experiences is a Cartesian mind leads to the stultifying consequence that self-attribution of experiences is impossible.<sup>21</sup>

But Kim errs, I think, at two points. First, Cartesian theories, both as Strawson represents them and as most if not all Cartesians represent them, do not purport to explain but rather presuppose the possibility of self-ascription of experiences. Secondly, Strawson maintains that under the assumption that subjects of experience are Cartesian egos, not just self-ascription of experiences, but any ascription of experiences, must be considered impossible.

## 4. The Thesis of Primitiveness

Strawson asserts that “the concept of a person,” in the technical sense in which he uses this term, is “logically primitive.” But the question what he means by this claim, and the question how he defends it, are matters of much dispute. He explains what he means in several passages, the chief of which reads as follows:

[1] . . . We are tempted to think of a person as a sort of compound of two kinds of subjects: a subject of experiences (a pure consciousness, an ego) on the one hand, and a subject of corporeal attributes on the other. . . . when we ask ourselves how we come to frame, to get a use for, the concept of this compound of two subjects, the picture—if we are honest and careful—is apt to change from the picture of two subjects to the picture of one subject and one non-subject. For it becomes impossible to see how we could come by the idea of different, distinguishable, identifiable subjects of experiences—different consciousnesses—if *this idea is thought of as logically primitive*, as a logical ingredient in the compound-idea of a person, the latter being composed of two subjects. For there could never be any question of assigning an experience, as such, to any subject other than oneself; and therefore never any question of ascribing it to a subject at all. So 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 analysed. It can exist only, if at all, as a secondary, non-primitive concept, which itself is to be explained, analysed, in terms of the concept of a person. . . .<sup>22</sup>

Farther down the page, he adds:

[2] The concept of a person is logically prior to that of an individual consciousness. The concept of a person is not to be analysed as that of an animated body or of an embodied anima. This is not to say that the concept of a pure individual consciousness might not have a logically secondary existence, if one thinks, or finds, it desirable. We speak of a dead person—a body—and in the same secondary way we might at least think of a disembodied person. A person is not an embodied ego, but an ego might be a disembodied person, retaining the logical benefit of individuality from having been a person.<sup>23</sup>

Farther on, he adds, by way of summary:

[3] . . . All I have said about the meaning of saying that [the concept of a person] is primitive is that it is not to be analysed in a certain way or ways. We are not, for example, to think of it as a secondary kind of entity in relation to two primary kinds, viz. a particular consciousness and a particular human body. . . .<sup>24</sup>

These passages indicate, I think, that Strawson assumes the following definition: ‘Concept *X* is logically primitive with respect to concept *Y*, or in other words, is logically

### Strawson's Argument against the Cartesian Thesis: An Interpretation

prior to concept *Y*, if (1) concept *X* cannot be derived from, be defined or explained by reference to, be constructed on the basis of, or otherwise be introduced into discourse by means of concept *Y*, while (2) concept *Y* can be introduced by means of concept *X*, and only by means of concept *X*. In passages [1]-[3], in which he explains what he means by claiming that the concept of a person is "logically primitive," he asserts several things about the connection between this concept and that of a pure subject of experience. In passage [1], he argues that the latter can be introduced into discourse only by means of the former, if it can be introduced at all, and in [2] and [3], he asserts that the former cannot be built up from the latter. But in passage [2], he asserts that the latter can be derived from the former. What he asserts about the connection between these two concepts, then, is this: the concept of a person satisfies both condition (1) and condition (2) of the definition just stated for being logically primitive with respect to that of a pure subject.

One might ask why Strawson affixes the qualifying term 'logically' to the terms 'primitive' and 'prior.' In a broad sense of the term, a question can be called a matter of logic, I would say, if in one way or another it has to do with the meaning or function of a linguistic expression. But the question whether one concept (e.g., that of a "person") is "primitive" in Strawson's sense with respect to another (e.g., that of a pure subject of experience) has to do with the ways in which the corresponding expressions (e.g., 'a person,' 'a pure subject of experience') are given meanings, and thus can be called a matter of logic.

Strawson's thesis of primitiveness, then, as I interpret it, can be stated as a conjunction of two distinct clauses: (1) the concept of a "person," in the technical sense Strawson gives this term, cannot be introduced into discourse by means of that of a pure subject of experience; and (2) the concept of a pure subject of experience can be introduced into discourse by means of that of a "person," and only by means of that of a "person." It is not especially difficult to see how he argues for these two clauses. As for clause (1), to say that the concept of a "person," in Strawson's technical sense, *can* be introduced into discourse by reference to that of a pure subject of experience, is to say that it can be introduced by reference to that of a Cartesian ego. But if his argument against the Cartesian thesis is sound, then the concept of a Cartesian ego is a pseudoconcept, and hence cannot be used to introduce the concept of a "person" into discourse. So his argument against the Cartesian thesis can be taken to be his argument for clause (1). As for the first part of clause (2), his position appears to be that the concept of a pure



subject of experience can be derived from that of a “person” by a process of abstraction—that is to say, by imagining the physical aspect of a “person” to fall away, which leaves only the mental aspect, or to put it another way, by leaving out of account the physical attributes, or to put it yet another way, the physical aspect, of a “person.” The second part of clause 2 is the assertion that there is no concept other than that of a subject of experience that has a physical aspect by means of which the concept of a subject of experience that has no physical aspect can be introduced into discourse. Strawson appears to take this to be obvious.

In the course of explaining his claim that the concept of a “person” is “logically primitive,” Strawson asserts that the concept of a disembodied person has a sound basis (see page 102 above, passage [2]). He elaborates on this at the end of his chapter “Persons”:

Earlier, when I was discussing the concept of a pure individual consciousness, I said that though it could not exist as a primary concept to be used in the explanation of the concept of a person . . . , yet it might have a logically secondary existence. Thus, from within our actual conceptual scheme, each of us can quite intelligibly conceive of his or her individual survival of bodily death. The effort of imagination is not even great. One has simply to think of oneself as having thoughts and memories as at present, visual and auditory experiences largely as at present, even, perhaps—though this involves certain complications—some quasi-tactual and organic sensations as at present, whilst (*a*) having no perceptions of a body related to one’s experience as one’s own body is, and (*b*) having no power of initiating changes in the physical condition of the world, such as one at present does with one’s hands, shoulders, feet and vocal chords. Condition (*a*) must be expanded by adding that no one else exhibits reactions indicating that he perceives a body at the point which one’s body would be occupying if one were seeing and hearing in an embodied state from the point from which one is seeing and hearing in a disembodied state. . . .<sup>25</sup>

Kim remarks:

. . . As Strawson . . . suggests, a disembodied mind can be referred to if it had some causal relations with a particular body in the past. . . .<sup>26</sup>

If I read this correctly, Kim certainly misses the point. For to speak of causal relations between two numerically distinct entities, a mind and a body, is to assume the truth of the Cartesian thesis. The position Strawson takes, which is contrary to the Cartesian thesis, is this: a subject of experience, an entity that has experiences, must be conceived, if it is to be conceived intelligibly, as an entity that has, or at any rate at one time had, physical attributes as well as experiences. The concession he makes in admitting the

intelligibility of the concept of a disembodied person is nothing more, and nothing less, than the concession represented by these words "or at one time had."

## 5. Objections and Replies

In reading over Strawson's statement of his argument against the Cartesian thesis, a number of objections occurred to me. I have, however, attempted to render each of these objections void by altering my interpretation of the argument in various ways. The result is that I am left with no objections that I consider to be damaging. I shall now, however, present and discuss several objections of the sort that have occurred to me. My purpose is twofold: to forestall these particular objections, and to develop my interpretation of Strawson's reasoning somewhat further.

### *Objection 1*

Strawson asserts that any particular that has a spatio-temporal position has a unique spatio-temporal position, i.e., one shared by no other particular.<sup>27</sup> In other words, he asserts that if a particular occupies a certain volume of space at a certain time, then no other particular occupies precisely that volume at that time. This principle figures in his account of the grounds of the possibility of identifying particulars, and in the present exposition, this account in turn figures in the defense of the third premiss of his argument against the Cartesian thesis (see pages 97 ff. above). But the principle is false. Thus, for example, two clouds of gas not reactive with each other moving through space on intersecting paths might occupy precisely the same volume of space at a certain point in time. The true principle, rather, is this: necessarily, at any given time, any particular that has a position in space has a unique position there, *or* is distinguishable in nonspatial terms from the entities that share its spatial position at that time.

I think that this objection is sound. It does not, however, serve to undermine Strawson's argument. Substituting the corrected version of the principle in question for the uncorrected version in the argumentation that I have attributed to Strawson necessitates no further adjustments in that argumentation. In particular, the thesis that the general system of relations among particulars, the system that encompasses all particulars, is at core a spatio-temporal system, and the thesis that Cartesian egos can be assigned positions neither within this spatio-temporal core nor anywhere else in the general system, remain standing.

### *Objection 2*

As Strawson himself appears willing to allow, it is intelligible to postulate that somewhere in the universe there is a society of thinking beings whose native perceptual and intellectual equipment differs in fundamental ways from ours, that these beings have

instituted a system of relations among particulars comparable, in its adequacy to the needs of a sophisticated culture, to the one that we use, that this system, unlike ours, is not at core a spatio-temporal one, and that it encompasses particulars of types that we do not recognize, including, notably, subjects of experience that do not have a corporeal aspect or any spatial aspect at any point in the duration of their existence. But such subjects are Cartesian egos, or if, besides experiences, nonspatial attributes of some sort pertain to them, are Cartesian egos in an extended sense of the term. Therefore, Strawson is mistaken to claim that the concept of a Cartesian ego cannot be given a sound basis.

My reply to this objection can be short. In arguing against the Cartesian thesis, Strawson's aim is to make several points concerning *human* thought and discourse. His claim that it is impossible to "frame" or "get a use for" the concept of a Cartesian ego is to be understood to mean that it is impossible for us human beings to incorporate this concept in our own thought and discourse, i.e., impossible for us to find any way to make statements about or to think about particular instances of the type that this concept represents.

It should be noted, however, that under the interpretation I have presented, Strawson's conclusions are to be taken to apply to human thought and discourse in general, and not to any system of thought or *weltanschauung* that might be said to pertain peculiarly to the English language or any other particular human language or group of languages. It should be noted that in defending his conclusions, he does not draw upon any peculiarity of English or any other language.

### *Objection 3*

Strawson maintains that there is no guarantee of a logical or conceptual sort that given a human body that behaves in a certain way, one and only one Cartesian ego is causally associated with it (see pages 97 ff. above). But he merely assumes, and does not establish, that in the realm of contingent fact no guarantee that this is so can be found.

My reply is this. Through his argument against the Cartesian thesis, Strawson attempts to establish that the concept of a Cartesian ego is unsound in that thoughts and statements about individual Cartesian egos are in principle impossible. Thus, no objection to this argument that merely presupposes and does not establish the soundness of this concept is acceptable. But the suggestion that a one-to-one correspondence might through observation of natural phenomena be discovered to obtain between human bodies and Cartesian egos presupposes the soundness of the concept of a Cartesian ego, for if this concept is not sound, then the question whether such a correspondence might obtain simply does not

arise. Therefore, this objection fails in that it presupposes, and does not establish, the soundness of the concept of a Cartesian ego. To put the point another way, the concept of a Cartesian ego must somehow be incorporated in language before the question of a one-to-one correspondence between human bodies and Cartesian egos can legitimately be raised, but in the objection just stated, there is no indication how this might be done.

*Objection 4*

Strawson claims that the following proposition is intelligible (though he does not claim it is true): When the body of a human being ceases to function, although the physical aspect of the individual in question falls away, as it were, the mental aspect continues to exist as all that is left of him or her, and functions as a subject of thoughts and memories, perceptual experiences, etc. He calls such a subject of experience, one that once had but no longer has a physical aspect, a disembodied person. He appears to take the position that a disembodied person in this sense of the term can be identified, at least by itself and to itself, as 'the pure subject of experience which at an earlier time was (Mr./Ms.) \_\_\_\_\_' (see pages 104-105 above). There is, however, no guarantee that at the moment the body of a human being ceases to function, many pure subjects will not spring into being, each boasting mental attributes characteristic of and memories from the life of the human being whose body it was. Hence, it cannot be considered legitimate to employ the definite description 'the pure subject of experience which at an earlier time was (Mr./Ms.) \_\_\_\_\_.' But there is no method besides the employment of this expression by which particular disembodied persons might be identified, and hence it is impossible in principle to identify them. Therefore, the concept of a disembodied person, defined as Strawson defines it, is in the final analysis unintelligible.

In reply, I grant that it is possible in principle that at the moment the body of a human being ceases to function, many pure subjects of experience might "spring into being," each having memories that stem from the experiences of the human being whose body it was. This is possible in that nothing about the relevant concepts, e.g., that of the body of a human being and that of a pure subject of experience, rules it out as a possibility. But let us suppose that following upon the death of a certain Mr. Jones, two disembodied persons in Strawson's technical sense, each with memories that spring from Jones's experiences, proceed to have qualitatively different series of experiences. One, for example, might perceive scenes in France, think to itself largely in Jones's halting French, and generally experience the optimism that was characteristic of Jones in his better moments, while the other might see only Idaho, think only in fluent English, and

generally experience the gloominess typical of Jones in times of difficulty. In a case of this sort, each subject might identify itself as the pure subject whose experience is continuous in certain respects with that of Jones and has been of such and such a sort since Jones's death. Now let us ask whether it is intelligible to postulate that following upon the death of a certain Mr. Smith, two numerically distinct pure subjects of experience might spring into being and proceed to have two qualitatively indistinguishable series of experiences which show certain continuities with Smith's experience. I think not. If the series of experiences in question are qualitatively indistinguishable, then what is there to distinguish them numerically—that is to say, in virtue of what are they to be regarded as two series and not one series? One naturally thinks of distinguishing them by reference to the subjects which have them: if one series pertains to one subject, and the other to a numerically distinct subject, then the two series are numerically distinct. But in virtue of what are the subjects in question to be regarded as two subjects and not one subject? The sum total of what can be said to identify the one subject—that since the death of Smith it has had experiences of this and that sort which show continuities in this and that respect with Smith's experience—is identical to the sum total of what can be said to identify the other. Hence, the subjects in question cannot be considered numerically distinct: they must be considered to be one subject. If, then, two numerically distinct disembodied persons represent continuations of the existence of the same human being, they have qualitatively different experiences. In order to save Strawson's thesis that the concept of a disembodied person is intelligible, then, it is only necessary to point out that a disembodied person could identify itself to itself as the pure subject whose experience has been of such and such a description since the death of (Mr. / Ms.) \_\_\_\_\_ and shows continuity in such and such respects with his or her experience.

#### *Objection 5*

Strawson's point about the logic of predicate-expressions is unsound. For it entails that if one can ascribe an experience of a certain type to oneself in that one has in one's vocabulary a predicate-expression bearing the appropriate meaning, then one can ascribe experiences of that type to others also in that one knows under what conditions it is correct to apply that predicate-expression to them (see pages 93-94 above). But this is false. For, clearly, one might distinguish sensations of a novel type in one's experience, sensations of a type one has never had or heard mention of before, and ascribe sensations of that type to oneself in words like "I'm having that funny feeling again," and yet never get around to working out criteria for ascribing sensations of that type to others.

This objection must, I think, be judged ineffective. Even if the possibility described is genuine, it does not follow that Strawson's point about the logic of predicate-expressions, or the first premiss of his argument against the Cartesian thesis, is unsound. For one need not deliberately work out criteria for the ascription of sensations of a certain type to others in order to be said to know under what conditions it is correct to ascribe sensations of that type to others. For it suffices that upon encountering a case in which those criteria are fulfilled, one recognize that it is correct to ascribe a sensation of the type in question.

Moreover, I think that a case can be made for saying that in a situation of the sort described, one would indeed have a conception of criteria for the ascription to others of sensations of the novel type. Suppose that a person *P* is in a position to ascribe sensations of a certain type *T* to himself, i.e., suppose that he possesses the intellectual resources requisite to saying with full comprehension and at times at which it is true, "I'm having a sensation of type *T* now." Then he has a more or less adequate conception of type *T*, for if he had no conception of type *T*, then, although he might perhaps have sensations of that type, in the way a dog might, he could not be said to be able to ascribe them at all. But if he has a more or less adequate conception of type *T* of sensations, then, clearly, he can recognize an accurate description of type *T*, though perhaps only with some difficulty, if he hears one. But this means that he is able to employ a criterion for the ascription of sensations of type *T* to others, for the following constitutes such a criterion: the uttering by another person of words that indicate the occurrence in him of a sensation of that type.

## 6. Summary

To repeat by way of summary, Strawson's argument against the Cartesian thesis reads as follows under my interpretation: 'If one can ascribe experiences to oneself in that one has the requisite predicate-expressions at one's command for use in speech and thought, then one can ascribe experiences to others also, in that one knows under what conditions it is correct to apply those predicate-expressions to others. But if one can ascribe experiences to others in that one knows under what conditions it is correct to apply the relevant predicate-expressions to them, then one can identify others in that a coherent account of how to go about identifying them can be formulated. But if subjects of experience are Cartesian egos, then no coherent account of how to go about identifying others can possibly be formulated. Therefore, if subjects of experience are Cartesian egos, then it is impossible in principle for one to identify other subjects, and impossible

in principle for one to ascribe experiences to other subjects or to oneself.' This argument is of the form 'if *A* then *B*, if *B* then *C*, if *D* then not-*C*, therefore if *D* then (not-*C* and (not-*B* and not-*A*)), and hence is formally valid. As for the first premiss, Strawson draws attention to the point, perhaps rather obvious in itself, that it is an essential characteristic of predicate-expressions that any given one can in principle be applied significantly to any number of individuals, and suggests that this principle entails this premiss. On page 94 above, I present a derivation of the latter from the former. As for the second premiss, as I state and argue for it, its defense lies in the thought that if no question at all arises of identifying subjects of experience that are numerically distinct from oneself, then there is no such thing as knowing under what conditions it is correct to ascribe an experience of this or that sort to another subject. See page 96 above for a slightly fuller statement. To defend the third premiss, I have drawn upon materials taken from Strawson's argumentation concerning the grounds of the possibility of identifying particulars. The basic thought in the defense of this premiss is that Cartesian egos cannot be assigned positions within the general system of relations among particulars. See pages 97-98 above.

I have attempted to find an interpretation of Strawson's argument against the Cartesian thesis under which that argument is sound and philosophically significant. Quite apart from the question whether I have succeeded, I think that Strawson deserves credit for demonstrating as forcefully as he has, with respect to the Cartesian thesis and a number of other topics, the considerable philosophical importance of the matter of the identification of particulars.

## NOTES

1. Peter F. Strawson, *Individuals: An Essay in Descriptive Metaphysics* (1959; rpt. Bristol, England: University Paperbacks, 1979), p. 100.
2. See *Individuals*, p. 15.
3. See *Individuals*, pp. 15-16.
4. Donald Sievert, "Strawson on Persons" (*The Modern Schoolman*, 48 (1971), 237-62), pp. 252-53.
5. R. L. Phillips, "Descriptive versus Revisionary Metaphysics and the Mind-Body Problem" (*Philosophy: Journal of the Royal Institute of Philosophy*, 42 (1967), 105-118), p. 109.
6. *Webster's Third New International Dictionary of the English Language* (unabridged), ed. Philip Babcock Gove (Springfield, Mass.: G. & C. Merriam, 1971).
7. *Individuals*, pp. 101-102.
8. *Ibid.*, p. 99.
9. *Ibid.*, p. 100.
10. *Ibid.*, p. 99n.

Strawson's Argument against the Cartesian Thesis: An Interpretation

11. *Ibid.*
12. See, for example:
  - H. D. Lewis, "Mind and Body—Some Observations on Mr. Strawson's Views" (*Proceedings of the Aristotelian Society*, 63 (1962–63), 1–22), p. 14.What is basically the same interpretation can also be found in these critics:
  - M. C. Bradley, "Mr. Strawson and Skepticism" (*Analysis*, 20 (1959), 14–19), p. 15.
  - David F. Pears, Part I of review of *Individuals* (*Philosophical Quarterly*, 11 (1961), 172–85), pp. 180–81.
  - R. L. Phillips, *op. cit.*, p. 109.
13. *Individuals*, p. 100.
14. *Ibid.*
15. *Ibid.*, pp. 25–26.
16. *Ibid.*, p. 101.
17. *Ibid.*
18. *Ibid.*
19. *Ibid.*
20. *Ibid.*, p. 6.
21. Chin-Tai Kim, "Cartesian Dualism and the Unity of a Mind" (*Mind*, 80 (1971), 337–53), p. 339.
22. *Individuals*, pp. 102–103.
23. *Ibid.*, p. 103.
24. *Ibid.*, pp. 104–105.
25. *Ibid.*, p. 115.
26. "Cartesian Dualism," p. 341.
27. *Individuals*, pp. 25–26.