GRAY, Paul Dennis, 1942-
A DEFENSE OF P.F. STRAWSION'S THEORY OF SELF.

The Ohio State University, Ph.D., 1970
Philosophy

University Microfilms, A XEROX Company, Ann Arbor, Michigan

Copyright by
Paul Dennis Gray
1971
A DEFENSE OF P. F. STRAWSON'S THEORY OF SELF

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Paul Dennis Gray, A.B.

* * * * * *

The Ohio State University
1970

Approved by

[Signature]
Adviser
Department of Philosophy
ACKNOWLEDGMENTS

I wish to thank all of the members of the Philosophy Department of The Ohio State University for their encouragement and instruction during my graduate study. And I wish to give special thanks to my advisor, Professor Bernard Rosen, for his time and thought, and, for his exemplary devotion to reason.
VITA

August 16, 1942 ....... Born - Salem, Ohio

1960 ............... Graduated, Salem Senior High School,
                 Salem, Ohio

1964 ............... A. B., Youngstown State University,
                 Youngstown, Ohio

1964-1969 .......... Teaching Assistant, The Ohio State
                 University, Columbus, Ohio

1969- ............... Instructor, Department of Philosophy,
                 Cornell College, Mt. Vernon, Iowa
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF SYMBOLS</td>
<td>vi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. THEORY PREFERENCE</td>
<td>4</td>
</tr>
<tr>
<td>Phenomena To Be Accommodated</td>
<td></td>
</tr>
<tr>
<td>Problems To Be Solved</td>
<td></td>
</tr>
<tr>
<td>General Extrinsic Theory Preference Principles</td>
<td></td>
</tr>
<tr>
<td>Defense of the Criteria</td>
<td></td>
</tr>
<tr>
<td>II. STRAWSON'S METAPHILOSOPHY</td>
<td>20</td>
</tr>
<tr>
<td>The Methods and Tasks of Philosophy</td>
<td></td>
</tr>
<tr>
<td>Critical Comments</td>
<td></td>
</tr>
<tr>
<td>III. DESCRIPTIVE METAPHYSICS AND CONCEPTUAL LAW</td>
<td>31</td>
</tr>
<tr>
<td>A Physical Law Model</td>
<td></td>
</tr>
<tr>
<td>IV. THE CARTESIAN THEORY</td>
<td>40</td>
</tr>
<tr>
<td>Description of the Theory</td>
<td></td>
</tr>
<tr>
<td>Strawson's Objections to the Theory</td>
<td></td>
</tr>
<tr>
<td>Critical Examination of the Theory</td>
<td></td>
</tr>
<tr>
<td>V. HUME'S THEORY</td>
<td>65</td>
</tr>
<tr>
<td>Description of the Theory</td>
<td></td>
</tr>
<tr>
<td>Strawson's Objections to the Theory</td>
<td></td>
</tr>
<tr>
<td>Critical Examination of the Theory</td>
<td></td>
</tr>
<tr>
<td>VI. THE NO-OWNERSHIP THEORY</td>
<td>83</td>
</tr>
<tr>
<td>Description of the Theory</td>
<td></td>
</tr>
<tr>
<td>Strawson's Objections to the Theory</td>
<td></td>
</tr>
<tr>
<td>Critical Examination of the Theory</td>
<td></td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>PH.1.</td>
<td>The Phenomenon of Agency</td>
</tr>
<tr>
<td>PH.2.</td>
<td>The Phenomenon of Re-Identification</td>
</tr>
<tr>
<td>PH.3.</td>
<td>The Phenomenon of Identification</td>
</tr>
<tr>
<td>PH.4.</td>
<td>The Phenomenon of Individuation</td>
</tr>
<tr>
<td>PH.5.</td>
<td>The Phenomenon of Ordinary Language</td>
</tr>
<tr>
<td>PR.1.</td>
<td>The Problem of the Nature of a Self</td>
</tr>
<tr>
<td>PR.2.</td>
<td>The Problem of Other Persons</td>
</tr>
<tr>
<td>PR.3.</td>
<td>The Problem of Solipsism</td>
</tr>
<tr>
<td>PR.4.</td>
<td>The Mind-Body Problem</td>
</tr>
<tr>
<td>PR.5.</td>
<td>The Problem of Personal Identity</td>
</tr>
<tr>
<td>P.1.</td>
<td>The Simplicity Principle</td>
</tr>
<tr>
<td>P.2.</td>
<td>The Consistency Principle</td>
</tr>
<tr>
<td>P.3.</td>
<td>The Ad Hoc Principle</td>
</tr>
<tr>
<td>CT</td>
<td>The Cartesian Theory</td>
</tr>
<tr>
<td>HT</td>
<td>Hume's Theory</td>
</tr>
<tr>
<td>NT</td>
<td>The No-Ownership Theory</td>
</tr>
<tr>
<td>ST</td>
<td>Strawson's Theory</td>
</tr>
</tbody>
</table>
INTRODUCTION

There is no doubt that in the last twenty years or so, what has come to be called Oxford or Ordinary-Language Philosophy has, in one form or another, been a dominant influence (perhaps the dominant influence) in Anglo-American philosophy. As well, a look at these last twenty years reveals a predominant interest in what has come to be called the Philosophy of Mind. Two philosophers who have figured most prominently in both the above "movement" and the above interest are Gilbert Ryle and Peter Strawson.

While Ryle concentrated upon an Oxford-type investigation of the "logic" of certain concepts (such as Will, Intelligence, Pleasure, and so on), Strawson has concentrated upon the broader concept (and topic) of the subject which has the Will, Intelligence, Pleasure, and so on. That is, Strawson has concentrated upon the concept of a self or of a person. What Strawson purports to have done is: (a) to have shown that a number of prominent theories of self fail, and (b) to have produced a theory of self which succeeds; that is, which is not subject to the objections Strawson raises against the other theories and, furthermore, which plays some role in the solution to or the dissolution of some related major problems in the philosophy of mind.

\[1\] The terms "self" and "person" will be used interchangeably throughout this study.
My primary goals in this dissertation will be, first, to develop and defend a context in which to judge any theory of self; that is, to try to discover and clarify the sorts of conditions a philosophical theory of self should be able to meet before it can be accepted as an adequate or as a preferable theory of self.

Secondly, I will carefully examine four theories of self (the Cartesian theory, Hume's theory, the No-Ownership theory, and Strawson's theory) with special attention to Strawson's critical remarks on the Cartesian, Humian, and No-Ownership theories, and with special attention to Strawson's own theory of self and its role in the solution to or dissolution of some major problems in the philosophy of mind.

Accordingly, in chapter i I shall be concerned with developing a plausible framework by which to judge the adequacy of any theory of self. Chapters ii and iii will contain a discussion of Strawson's metaphilosophy. Chapter iv consists of a critical examination of Strawson's version of the Cartesian theory and Strawson's objections to it. In chapter v the theory of self of David Hume is examined. Chapter vi deals with a theory, called by Strawson the No-Ownership theory, which is purported to be found in the writings of Ludwig Wittgenstein and Moritz Schlick. It, too, is carefully explicated and then critically examined. In chapters vii, viii, ix, and x Strawson's theory is subjected to a very thorough critical examination. I attempt to discover how and how well Strawson's theory can better accommodate a range of phenomena (better, that is, than the other theories of self) and how and how well Strawson's theory can
solve or dissolve a set of related problems in the philosophy of mind. In the concluding chapter (chapter xi) I offer some suggestions toward a theory of self preferable even to Strawson's (some suggestions, that is, toward resolving some of the problems that arise even with Strawson's theory). Finally, I reflect upon the value and the limitations of this study in the philosophy of self.
Like many philosophical problems, the problem of self is best seen as primarily the problem of producing a theory of self which will best accommodate a range of relevant phenomena and which will solve more relevant problems and generate fewer objections than alternative theories of self. Let us look for a moment at the problem of free-will. Some of the phenomena which a theory of free-will must accommodate are: (1) the feeling of freedom, (2) the common belief that all events have a cause, (3) the common appropriate use of phrases like: a. "John did it, but not freely", b. "Jack was free to choose it or not", and c. "I did x freely and thus I am responsible for it", and so on.

By phenomenon I mean a prima facie datum. It is not the conclusion to a sound argument. Nor is it a result derived from a correct theory. It is rather, what appears to be an obvious truth or fact. And it is to be treated as a truth or fact, unless a theorist can and does offer reason(s) to show that it is not justifiably treated as a fact. Put somewhat differently, a phenomenon is to be treated as a fact, unless it is shown to be illusory (that is, shown to only appear to be true, though it is not true).
A theorist, then, can accommodate a phenomenon either by making his theory consistent with the phenomenon or by showing that the phenomenon is illusory. If it is illusory, then, of course, the theorist need not make his theory consistent with this phenomenon.

Another test of a theory of free-will (besides adequacy of phenomena-accommodation) is its ability to solve certain relevant problems and its success in generating few (and preferably, no) theoretical objections. For example, a free-will theorist must solve the problem of the reconciling our conflicting common-sense beliefs on this issue. It is common-sense to believe that persons do have free-will; yet it is also common-sense to believe that all events, actions, ...(of persons) are caused. If a free-will theory fails to reconcile these beliefs (solve this problem), this is a mark against that theory. That is, this failure would constitute an objection to the theory.

There is still a third test for theory-success. A theory is to be preferred if it has certain extrinsic features; such as, simplicity (in a sense to be specified later), consistency with other well-established theories, and so on. For example, a free-will theorist would have to concern himself with the question of whether or not his developed theory is consistent or inconsistent with the well-established relevant theories in science. In the case of a determinist, the problem of the relevance and consistency of his theory and quantum theory would have to be solved.

The above, then, is a short example of the testing of a theory of free-will. What I shall do in this dissertation is, first, develop
a plausible structure for the judging of the preferability, the success or failure, of a theory of self; and secondly, judge a set of theories of self in terms of this structure, with the person-theory of Peter Strawson as the focal point.

THE STRUCTURE

I will now set out as clearly and as cogently as I can what I take to be the relevant phenomena of the problem of self, the relevant problems to be solved or dissolved, and the general extrinsic theory preference principles to be applied.

I. PHENOMENA TO BE ACCOMMODATED

Before listing the phenomena which a theory of self should accommodate, I shall try to further clarify the concept of a phenomenon and especially its relationship with what are called "facts". Broadly speaking, the major requirement (though not the only requirement) that a theory should satisfy is that it be consistent with the facts. Now, facts have at least two sources: (1) Facts are the results derivative from a correct theory. For example, a number of facts are derivable from Einstein's Theory of Relativity (certain predictions, for example). (2) Facts are the conclusions to sound arguments. For example, if the standard argument,

1. All men are mortal.
2. Socrates is a man.
Therefore, Socrates is mortal.

is sound, then, "Socrates is mortal" is a truth or fact.
The above are clear examples of facts. But besides facts, which are clearly relevant to the adequacy of a theory, there are what I shall call phenomena. A phenomenon is not a fact. But, a phenomenon is a *prima facie* datum, and it is justifiably treated as a fact or as a truth, unless it is shown to be illusory. For example, it seems obvious that the earth is relatively flat. That the earth is relatively flat is a phenomenon. But, in the context of contemporary physical theory (well-established) and of a number of observations made by scientists, this phenomenon has been shown to be illusory. The earth is, in fact, (a fact in relation to a theory) relatively round. Thus, one is not justified in treating the earth's relative flatness as a fact. But it still is a phenomenon (though an illusory one).

On the other hand, it is a phenomenon that there are many persons in this world. That there is more than one person appears quite obviously true. But, perhaps, this is just an illusory phenomenon. It is logically and physically possible that there is now but one person in the world. Now, if a theorist can give a good reason or argument (or more than one, of course) or if a theorist can produce a well-established theory, that shows that there is but one person (or none) in the world, then, that there are a number of persons in the world, will be an illusory phenomenon. And that there are a number of persons will no longer be justifiably treated as a fact. On the other hand, if no such reasons, arguments, or theory can be produced, then one is justified in treating the existence of many persons as a fact.
Some miscellaneous examples of (as far as we now know) non-illusory phenomena would be:

1. Horses need food in order to survive.
2. Some men enjoy fishing.
3. There are birds.
4. Richard Nixon is the present President of the United States.

In short, then, I wish to hold that a theory should, to be justifiably considered a good theory, be consistent with the relevant facts, and with the relevant phenomena which are justifiably treated as facts. To repeat, a theory must not ignore its relevant phenomena. But, it can accommodate it in one of two ways. Either it must show that the phenomenon is illusory, in which case the phenomena is not justifiably treated as a fact. Or it must be consistent with the relevant phenomenon, which if non-illusory is justifiably treated as a fact, which the theory must be consistent with.

Now, in light of the above, I will list what I believe to be the phenomena relevant to persons and the development of a satisfactory theory of self. The following are such relevant phenomena:

PH.1. The Phenomenon of Agency: Persons are agents. That is persons do things and things are done to them.

PH.2. The Phenomenon of Re-identification: There are clear, ordinary, and generally successful (though not infallible) ways of re-identifying persons. That is, there are adequate ways of telling that a person whom you observe (see, hear, ...) is the same person you observed before. For example, there are clear, ordinary, and generally successful ways of telling that the woman now in your house (your wife) is the same woman (your wife) whom you married years ago.
The Phenomenon of Identification: There are clear, ordinary, and generally successful (though not infallible) ways of identifying persons. That is, there are adequate ways of telling that a person whom you observe (see, hear, ...) is Jack, or John, or Jill, or whomever. For example, you know who your wife is. You can pick her out of a crowd. You can correctly judge that she is your wife. In short, you can identify her. The major difference between re-identification and identification is that with the former (re-identification), one is judging that a person P is the same person as a person whom you have observed before; in the latter (identification), one is judging that a person P is person P rather than person Q or person R or person S, and so on.

The Phenomenon of Individuation: There are clear, ordinary, and generally successful (though not infallible) ways of individuating persons. That is, there are adequate ways of telling that in a given context there is one person, or two, or three, or more. For example, there are clear, ordinary, and generally successful ways of telling that one has two children, not just one; that one is in a group of six persons; and so on. The difference between, for example, identifying and individuating persons is that in the former (identification) one is interested in the question, "Which person in the group is (for example) Jack?" Whereas in the latter (individuation), one in interested in the question, "How many persons are there in the group of which Jack is a member?"}

1I should mention that I am not now interested in the close relationships between re-identification, identification, and individuation. That there are such relationships, I willingly concede.
PH.5. The Phenomenon of Ordinary Language: The theory should be able to accommodate any ordinary, meaningful, obviously true statement involving persons. Now, there are, of course, many such statements. I shall list below merely a sample (using, of course, fictitious names that can easily be replaced by non-fictitious names such as to render the statements obviously true):

1. "John is now thinking."
2. "Jack is walking along the street."
3. "Jill hopes she will be able to come."
4. "Bob died."
5. "Bill believes in God."
6. "There are a number of persons over there."
7. "That is Joan."
8. "John is building a model plane."

It should be observed that the failure of a theory of self to be able to grant that persons can be identified, will thereby also be its failure to accommodate 7. above. For, surely, it cannot be obviously true that "That is Joan" if persons cannot be identified. (The theorist could, of course, try to show that it is illusory that "That is Joan" is obviously true.) Likewise, the failure of a theory to be able to accommodate individuation of persons will thereby be its failure to accommodate 6. above. For, surely, it cannot be obviously true that "There are a number of persons over there" if persons cannot be individuated.

I mention the above because I wish (again) to concede that there are relationships between the phenomena listed. It is even the case that a theory's failure to accommodate one phenomenon may entail its failure to accommodate another phenomenon. But I do not see this as problematic. And, I do wish to stress that the phenomena I have listed are different; and, thus, are justifiably distinguished.
II. PROBLEMS TO BE SOLVED

I shall now list what I believe to be the major problems that a theory of self must solve or dissolve. Immediately following the list, I shall make some comments upon it.

PR.1. The Problem of the Nature of a Self or Person: The theory must clearly specify what sort of entity a person is. \(^2\) For example, is a person a physical object, a mental substance, a process, a series of mental events, and so on?

PR.2. The Problem of Other Persons: The theory must be able to explain our knowledge of the existence of other persons. That is, it must be able to grant that we do know that there are other persons and it must explain how we can know this.

PR.3. The Problem of Solipsism: The theory must be able to explain our knowledge of things other than ourselves. It must be able to grant that we do know that there exists something other than ourselves and it must explain how we can know this.

PR.4. The Mind-Body Problem: The theory must clarify the differences between the physical aspects of a person and the mental aspects of a person. And the theory must explain the relationship between these two aspects. This problem can be seen as a subdivision of PR.1. But, this problem is of such importance that it deserves a special listing. Also, it historically is usually treated as a separate problem.

---

\(^2\) I use the term "entity" as an ontologically neutral term. That is, it could refer to a substance or to a non-substance; to a process or to a non-process, or whatever.
PR.5. The Problem of Personal Identity: The theory must explain how a person remains the same person (remains a person) despite the fact that he changes. It must explain how we can know that a person is the same person.

Now, for some comments on the above list.

C.1. That a theory of self needs to clearly specify what a self or person is (PR.1.) surely needs no defense.

C.2. It should be mentioned that problems PR.2. through PR.5. depend on the existence of phenomena. That is, each problem is a problem only because certain phenomena do exist. For example, the phenomenon involved in The Problem of Other Persons is the existence of our knowledge of other persons. The phenomenon involved in The Problem of Solipsism is our knowledge of other things. And so on.

C.3. Given C.2., one legitimate way for a theory to solve one or more of these problems is to show that the phenomenon upon which the problem depends, is illusory. For example, if a theorist can show that it is illusory (and thus not justifiably treated as a fact) that we know that there are other persons, then, of course, the theory no longer needs to solve The Problem of Other Persons, for there is no knowledge of other persons which must be explained. The same applies to PR.3. through PR.5.

It follows from the above that there are phenomena which a theory of self should accommodate other than those listed in I. (PHENOMENA TO BE ACCOMMODATED). These are those that have been traditionally bound up with the problems listed in II. (PROBLEMS TO BE SOLVED). I have left them so bound.
Lastly, it should be observed that if there is a serious objection to the theory which happens not to fall into the above categories (that is, into PR.1-5) then even that objection (which, of course, would constitute an additional problem) would have to be effectively resolved. If there is a failure to resolve it, then this should be an added factor in deciding the best theory among a set of competing theories. The more unresolved objections (unsolved problems), the less preferable the theory.

III. GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES

P.1. The Simplicity Principle: A theory $T_1$ is preferable to a theory $T_2$ if $T_1$ is simpler than $T_2$. The problem of a careful and defensible criterion for theory-simplicity is a topic of much contemporary research and controversy. However, there are some relatively uncontroversial marks of simplicity. For example, A theory $T_1$ is simpler than a theory $T_2$ if $T_1$ requires the recognition of fewer types of entities than $T_2$.

P.2. The Consistency Principle: A theory $T_1$ is preferable to a theory $T_2$ if $T_1$ is consistent with other well-established theories. If, for example, there are any well-established scientific theories relevant to the problem of self, a theory of self would have to show itself consistent with the data (facts) derivable from these theories.

P.3. The Ad Hoc Principle: A theory $T_1$ is preferable to a theory $T_2$ if $T_1$ uses fewer ad hoc devices than $T_2$. By an ad hoc device

---

3The consistency must not simply result from ad hoc additions. See forthcoming P.3.
device is meant an added explanation or revision which is introduced solely to enable the theory to escape from evidence or argument which is detrimental to the theory, and which does not follow from the theory, and is not a natural or obvious consequence of the theory.

The above, then, is the context in which I shall examine and judge competing theories of self. However, before proceeding, it should be shown that this set of preference-criteria which I have developed is not arbitrary. It is, in fact, quite in conformity with generally accepted criteria for theory-preference in the sciences.¹

In relation to I. (PHENOMENA TO BE ACCOMMODATED) it is quite clear that any scientific theory has a range of phenomena which it must accommodate. In the case of the Copernican theory, the accommodation must be of a plethora of astronomical phenomena. In the case of the Theory of Evolution, the phenomena to be accommodated are largely biological and geological phenomena. And so on.

In relation to II. (PROBLEMS TO BE SOLVED) it is, again, clear that one of the tasks of a scientific theory is to solve certain problems. For example, of the Copernican Theory as developed and defended in De Revolutionibus Thomas Kuhn writes, "The De Revolutionibus was written to solve the problem of the planets, which, ¹

¹I am not claiming to be using all of the generally accepted criteria for theory preference in the sciences. I am merely claiming that those criteria that I am using are generally accepted criteria for theory preference in the sciences.
Copernicus felt, Ptolemy and his successors had left unsolved." The "problem of the planets" was very largely the problem of precise specification of the positions of the planets at any given time. Likewise, says Kuhn, "Conceptions like Bohr's atom [Bohr's Theory of the Atom] and Einstein's finite but unbounded space [Einstein's Theory of Relativity] were introduced to solve pressing problems in a single scientific specialty."

It should be added that in terms still of II. (PROBLEMS TO BE SOLVED) but of II. put negatively (that is, that a theory must be able to resolve serious objections) it is a recognized part of the defense of a scientific theory that the theory be able to meet serious objections. For example, Psycho-Analytic Theory must deal with the serious objection that the theory is "unscientific". More specifically, it must somehow resolve the objection that its use of such terms as "unconscious forces", the "id", and so on, are terms which have no empirical content and thus does not enable statements of the theory in which these terms occur to be empirically confirmed or disconfirmed. In so far as the theory fails to meet this serious objection, it is less preferable than a competing theory (all else being approximately equal) which is not subject to this objection.

Finally, in relation to III. (GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES) again I can show that these Principles are

---


6Ibid., p. 229.
quite generally accepted principles of preference for scientific theories. For example, concerning P.1. (The Simplicity Principle) says Philipp Frank, "the requirements for the acceptance of a theory by scientists in the modern sense are 'the agreement with observation' and 'simplicity'." Furthermore, Carl Hempel writes, "Another aspect that affects the acceptability of a hypothesis is its simplicity, compared with that of alternative hypotheses that would account for the same phenomena." He continues, "... if two hypotheses accord with the same data and do not differ in other respects relevant to their confirmation, the simpler one will count as more acceptable."  

In support of P.2. (The Consistency Principle) I quote Hempel who writes, "the credibility of a hypothesis will be adversely affected if it conflicts with hypotheses or theories that are accepted at the time as well-confirmed."  

Lastly, a little history of science can be used to support P.3. (The Ad Hoc Principle). One of the reasons for the abandonment of the Phlogiston theory of combustion was that in order for the theory to accommodate all the phenomena, some most implausible assumptions had to be added to the theory. In particular, the


9Hempel, Philosophy of Natural Science, p. 39.
theory claimed that when a metal was heated, the result would be a Calx (what one would now call an oxide compound) and Phlogiston. According to the theory, if one heats a metal in a closed space, the metal "gives off" Phlogiston to the air and when the air becomes saturated with Phlogiston, combustion ceases (since the air can no longer "receive" the Phlogiston). If one opens the container, then, of course, more air is available (air that is not saturated with Phlogiston) and thus combustion can continue. The Phlogiston was postulated, then, to account for certain phenomena (the alteration of the metal and the air, the cessation of combustion, and so on). But then another relevant phenomenon was discovered; namely, that the weight of the Calx (formed, according to the theory, by "giving off" Phlogiston) was more (not less) than the weight of the original metal (before combustion)! To accommodate this phenomenon, the Phlogiston theorists added the assumption that the Phlogiston had a negative weight! And as more "embarrassing" phenomena became known more and more assumptions of this sort had to be added.

On the other hand, a competing theory (the Oxygen theory of combustion) developed by Antoine Lavoisier, did not require such ad hoc devices and (for this reason as well as others) it is the latter theory which is today preferred.  

---

10 For further evidence that science has used and does use P.3. as a theory-preference principle, see James Conant's, Science and Common Sense (New Haven, Yale University Press, 1951), especially pages 181 to 193; and, more recently, Leonard Nash's, The Nature of Natural Science (Boston, Little, Brown, 1963), especially pages 260 to 284.
So far, I have explained and defended the particular approach to "doing philosophy" (that is, to solving certain philosophical problems) which I shall use in the remainder of this study. It might be wondered if there are any precedents for this approach. There are. Though I do not wish to suggest that this approach is identical with the approaches of the well-known philosophers which I will soon mention, I do wish to suggest that the approach I shall use is in some ways similar to theirs.

For example, in chapter viii of William James' *Some Problems of Philosophy* (1911) James decides between two theories (Monism and Pluralism) on the basis of such considerations as, the ability of the theory to be consistent with common sense and science, the ability of the theory to meet serious objections brought against it, the special advantages of the theory, and so on. As well, though John Dewey does not set out the somewhat elaborate structure that I have, nonetheless, his approach, like the one I am using, holds that "philosophy has no private store of knowledge or of methods for attaining truth" and that philosophy should, broadly speaking, use common sense and science and especially scientific methods to solve philosophical problems.11

And lastly, I quote C. J. Ducasse who writes,

The theories that constitute a philosophy--whether a philosophy of religion, of art, of morality, or of anything else--are of course speculative inventions; but they are inventions that purport to be, or to be capable of providing, true answers to certain questions, and whose claim to be therefore discoveries is to be tested,

like other claims of truth, not only in respect to internal coherence, but also by reference to the relevant sort of facts.

But in this respect the case of philosophical theories, and that of physical, or biological, or mathematical or other scientific theories, is exactly alike.\textsuperscript{12}

CHAPTER II

STRAWSON'S METAPHILOSOPHY

Before examining the Cartesian theory of self (the first of four theories of self to be examined) I wish to examine the general conception of philosophy of Peter Strawson. The reason (justification) for this preliminary examination is that in one way or another (either via his critical arguments on the theories of self I shall examine or via his arguments in favor of his own theory of self) Strawson's views are the focal point of this study and to understand his arguments and their defense it is often essential to understand his view concerning the goals of philosophy and the methods appropriate to philosophy (especially, metaphysics).

For Strawson, philosophy is analysis. But not analysis as understood by Russell, or by Moore, or by Carnap, and others. Rather, Strawson's understanding of analysis derives largely from the later Wittgenstein.

There are at least two things I wish to say about Wittgenstein's conception of philosophy as exposit in *Philosophical Investigations*: 1. Philosophy is conceived of as a method of dissolving philosophic problems by a careful description of the uses made of certain key words or concepts relevant to the particular philosophic problem.
Philosophy is a therapeutic; a way of curing philosophic diseases called philosophic problems. 2. Philosophy is also conceptual clarification (whether or not that clarification results in the solution to a philosophic problem). This clarification results from careful description of the uses of the relevant words or concepts. So, for the later Wittgenstein, Philosophy is--at least, 1. conceptual therapy and 2. conceptual clarification.

Strawson agrees that 1. and 2. are a part of the tasks of philosophy. Indeed, Strawson calls 1. Conceptual therapy. It is, says he, the resolving of philosophic paradoxes; the clearing up of conceptual confusion; the diagnosing of philosophic disorders.

It is correcting conceptual distortions. And the method is "that of coming to understand philosophically puzzling concepts by carefully and accurately noting the ways in which the related linguistic expressions are actually used in discourse." But Strawson also agrees that philosophy involves more than just solving (or dissolving) philosophic problems (paradoxes; puzzles). It also involves "conceptual anatomy". This is the description of the "logic" of certain terms, for its own sake. By doing this, however, we would presumably gain a better understanding "... of the concepts and categories in terms of which we carry on our thinking ...".

2 Ibid., p. 104.
3 Ibid., pp. 103, 106.
Tasks 1. and 2. Strawson describes as the analytical tasks of philosophy.

But, Strawson holds that there are other tasks, as well. And here, I think, Strawson is going beyond Wittgenstein and other so-called "Linguistic Philosophers". These further tasks of the philosopher bring into play what Strawson calls philosophic imagination. They (3. and 4.) represent the imaginative side of philosophy. The third task Strawson calls the explanatory task. The fourth Strawson calls the inventive or constructive task. (Let us designate tasks 1., 2., 3., and 4. T.1., T.2., T.3., and T.4.) T.3. involves going beyond the description of how our concepts work and deals with why they work. That is, it attempts to discover and explicate why we have and use the concepts we do have and use. The assumption here is that we have certain concepts because we are of a certain nature ourselves, and, because the world is a certain way (has a certain nature). The method used to discover the above is as follows: "to imagine our experience being different in fundamental ways, and then to consider how our conceptual apparatus might naturally be adjusted to accommodate these differences. In seeing this, we also see how our concepts as they are, are rooted in the world as it is."^5

An example of T.3. would presumably be (though Strawson does not assert this) positing a No-Space world (an auditory world, a world which is non-spatial; a world only of sound(s)) and then trying


^5 Ibid., p. 107.
to discover if certain concepts which are available to us now (in our spatio-temporal world) would or would not be available in the No-Space World. One could ask, for example, could there be a concept of oneself (in contrast to that which is not-one self) in this Auditory World?

T.lj. also involves the philosopher's imagination; but involves it in a somewhat different manner. To engage in T.lj. is "... to consider how, without the nature of the world being fundamentally different, we might nevertheless view it through the medium of a different conceptual apparatus, might conduct our discourse about it in forms different from, though related to, those which we actually use." Strawson again offers no examples, but an example of T.lj. might be Strawson's consideration (in chapter iv of Individuals) of a conceptual scheme which dispenses with the use of demonstratives. The trouble is that here—though Strawson is working with a modification of our conceptual scheme, he is also working with a view of the nature of the world which is very different from an ordinary view; namely, the view that the world consists of monads. Thus, Strawson simply does not give any historical example of the actual occurrence of T.lj.

6This No-Space world is discussed at length in chapter ii of Strawson's book, Individuals: An Essay in Descriptive Metaphysics (New York, Doubleday and Co., 1959); hereinafter referred to as, Individuals.

In a later metaphilosophical essay ("Carnap's Views on Constructed Systems Versus Natural Languages in Analytic Philosophy")8 Strawson again offers no example of T.h. but does say that "Some metaphysics is best, or most charitably, seen as . . ." T.h. But, he further suggests that these modified conceptual schemes are most often presented not as modifications of our ordinary conceptual schemes, but rather, are presented as explications of our ordinary conceptual scheme.

In Individuals Strawson restricts his metaphilosophical discussion to Metaphysics. There are, he states, two types of Metaphysics: 1. descriptive and 2. revisionary. A philosopher engaging in descriptive metaphysics attempts to accurately describe the actual conceptual scheme that we use in thinking about ourselves and the world. More accurately, the descriptive metaphysician aims to "lay bare the most general features of our conceptual structure . . ."9 Strawson offers Aristotle and Kant as historical examples of descriptive metaphysicians. On the other hand, revisionary metaphysics goes beyond description; for it is the attempt to produce a better conceptual scheme (better than the one we actually possess). As historical examples of revisionary metaphysicians, Strawson lists Descartes, Leibniz, and Berkeley.

---


9 Individuals, p. xiii.
It might be wondered, which of the four tasks of philosophy descriptive and revisionary metaphysics involve. Strawson does not answer this question for us. But it would seem that descriptive metaphysics would most clearly involve T.2. (description of our actual use(s) of various concepts) and T.3. (the attempt to discover and describe why we have the concepts we have, involving, perhaps the imagining of "worlds" different from our own). But, it could even involve T.1. and T.4., in that doing T.1. and T.4. could perhaps enable us to better understand our actual conceptual scheme. That is, in so far as resolving conceptual confusions, puzzles, ...(T.1) and modifying our conceptual scheme and seeing the results (T.4.) does enlighten us about our own conceptual scheme, then T.1. and T.4. could very well be part of doing descriptive metaphysics.

On the other hand, it is not at all so clear which of the tasks would be involved in revisionary metaphysics, since it is not at all clear what it is that the revisionary metaphysician is doing. We are told that the revisionary metaphysician is attempting to make a better conceptual scheme. But what does this mean? Does it mean, one possessing more clarity? less distortion, less paradox? Does it mean one which is simpler? Strawson does not tell us. But on the assumption that to produce a "better" conceptual scheme is to do any

Evidence that descriptive metaphysics would involve T.3. as well would be that Strawson claims to be doing descriptive metaphysics in Individuals, and chapter ii involves the imagining of a No-Space world.
or all of the above, then again it would seem that any or all of the four tasks of philosophy could play a role in producing this "improved" conceptual structure. In short, then, doing metaphysics (either descriptive or revisionary) could consistently call into play any or all of the four general tasks of philosophy.

We now have a relatively clear grasp of the general tasks and methods of philosophy, as considered by Strawson. But, let us now look at two very intriguing quotes from *Individuals*.

1. Up to a point the reliance upon a close examination of the actual use of words is the best, and indeed the only sure, way in philosophy. But the discriminations we can make, and the connexions we can establish, in this way, are not general enough and not far-reaching enough to meet the full metaphysical demand for understanding. For when we ask how we use this or that expression, our answers, however revealing at a certain level, are apt to assume, and not to expose, those general elements of structure which the metaphysician wants revealed. The structure he seeks does not readily display itself on the surface of language, but lies submerged. He must abandon his only sure guide when the guide cannot take him as far as he wishes to go.\(^{11}\)

2. Metaphysics has a long and distinguished history, and it is consequently unlikely that there are any new truths to be discovered in descriptive metaphysics. But this does not mean that the task of descriptive metaphysics has been, or can be, done once for all. It has constantly to be done over again. If there are no new truths to be discovered, there are old truths to be rediscovered. For though the central subject-matter of descriptive metaphysics does not change, the critical and analytical idiom of philosophy changes constantly. Permanent relationships are described in an impermanent idiom, which reflects both the age's climate and the individual philosopher's personal style of thinking.\(^{12}\)

\(^{11}\) *Individuals*, pp. xiii, xiv.

But, what is this "... structure which the metaphysician wants revealed"? And what are these "old truths"? And, what could such "permanent relationships" be? I shall devote the next chapter to suggesting some answers to these questions. But before that, I will offer some general comments on Strawson's conception of philosophy as we now understand it.

CRITICAL COMMENTS:

C.1. The major weakness in Strawson's description of the method of Philosophy is its neglect of the non-linguistic "given" (phenomena) which surely must be considered relevant to the solution to at least some philosophic problems. For example, it seems ridiculous to suppose that the problem of the existence of God can be solved or dissolved simply by linguistic inquiry. Many of the arguments that purport to establish the existence of God rely upon purported factual claims about the world. In order to defend or refute these arguments, it is obviously relevant and even necessary to go beyond linguistic or conceptual data to other factual or phenomenal data. Strawson's method, as described by Strawson, reveals a striking neglect of this fact. But interestingly enough, while Strawson's description of his method neglects this (non-linguistic factual and phenomenal data) his actual use of his method reveals (as I shall later show) a clear reliance upon and appeal to facts and phenomena. For example, in his refutations of the Cartesian Theory of Self (which we shall soon examine) Strawson relies heavily upon the inability of CT to identify or re-identify or individuate a Cartesian person. From the
inability of CT to enable persons to identify, etc. persons, Strawson concludes that CT is not an acceptable theory of self.

Indeed, it is, I think, worthwhile to point out that Strawson's method of doing philosophy (as opposed to his description of the method) is very traditional. It is simply not the case that he engages only in the investigation and discussion of ordinary language use of words relevant to philosophic problems; whether in relation to T.1., T.2., T.3., or T.U. Indeed, Strawson does very very little of such linguistic or conceptual analysis or elucidation. For example, with respect to the term "person", Strawson does not ever once try to trace ordinary language uses of this term. Furthermore, Strawson uses argumentation, and generally, traditional forms of such argumentation. For example, in one of his arguments, he tries to show that a theory is unacceptable because it is forced to be inconsistent. In another, he attempts to show that a theory is forced to deny a prima facie fact. And so on.

C.2. It seems to me that Strawson's claim that descriptive metaphysics does not discover new truths is very misleading. In so far as, by doing descriptive metaphysics, one can generate arguments for and arguments against certain philosophical claims or theories (as Strawson does), one surely is involved in the discovery of new

---

\[^13\] Individuals, p. 92.

\[^{14}\] Ibid.
truths. And surely, in so far as Strawson's Individuals is an example of what Strawson calls descriptive metaphysics, it involves the critique of and the defense of philosophic theories. Furthermore, presumably Strawson would be willing to assert that the CT is incorrect or inadequate and his theory is correct or adequate. If so, this again is going beyond mere re-description of old truths. It is, is it not, the age-old attempt to discover (the) truth(s) about the world; the correct view of the way the world is; via a defensible theory.

In short, in spite of what some critics think and Strawson claims, Strawson is, at least with respect to the problem of self, engaging in the most traditional philosophical methods and goals; he attempts to refute claims and theories by clarifying them and arguing against them; and he attempts to defend claims and theories by clarifying them and arguing for them; and why(?), to arrive at the correct (most defensible) claims and/or theories?

C.3. Lastly, it should be observed that there is nothing incompatible about conceptual analysis (and descriptive metaphysics) as understood by Strawson and the method of philosophy which I am using and defending in this study. In so far as conceptual elucidation (analysis)(and doing descriptive metaphysics) provides for

---

general clarity, in so far as it can help solve or dissolve certain philosophical problems, it is to be desired and promoted. What one should not do however, is suppose that all philosophical problems are able to be solved or dissolved merely by linguistic (conceptual) description, explanation, or investigation. Specifically, the problem of self cannot thus be solved. Theory creation, theory critique, argumentation and appeal to phenomena are also necessary.
I wish now to examine the conception of descriptive metaphysics in more detail. Once again, I am discussing these issues not because they are intrinsically interesting (as they are), but because Strawson will be making descriptive metaphysical claims, sometimes in the form of premises to his arguments. So, it is imperative both in order to understand his arguments and in order to judge them, that one is as clear as possible about the epistemological status of descriptive metaphysical claims.

I shall begin by examining some examples of what would seem to be claims derived from descriptions of our conceptual scheme. The following then is a partial list of descriptive metaphysical statements:

a. If our conceptual scheme is one of a single spatio-temporal system of material things, then it is necessary that we are sometimes able to re-identify non-continuous particulars in this scheme.

b. "all deaths are necessarily deaths of creatures." OR, necessarily, the concept of death is not intelligible without the concept of a creature.

c. It is necessary that if one ascribes states of consciousness to oneself, one can also ascribe or be prepared to ascribe states of consciousness to others.

d. It is necessary that "one can ascribe states of consciousness to others, only if one can identify these others."
e.1. In order to identify persons, persons must be such that both states of consciousness and corporeal characteristics are ascribable.

e.2. "a person necessarily has corporeal attributes, as well as other kinds of attributes."

f. If x is a P-predicate (a particular sort of predicate to be explained later), then it is essential (necessary) that x is self-ascribable otherwise than on the basis of observation of the subject of them, and other-ascribable on the basis of behavioral criteria.

Some crucial questions arise now concerning the modal term "necessary" (necessarily, must, ...) which occurs in or beside each of the descriptive metaphysical statements; for example, (1) "What sort of necessity is it (logical, practical, psychological, ...)?" (2) "How are such necessities discovered (observation, experiment, intuition, induction, ...)"?

Strawson nowhere tries to deal with these questions. Yet, clearly, a large part of his philosophical success depends upon relatively plausible answers to the above questions. I shall now try to fit Strawson's views concerning necessity into a relatively familiar framework and thereby make his views intelligible and, indeed, at least as plausible as that framework. ²

It is now widely accepted that Laws of Nature (Physical Laws) are descriptive and not prescriptive statements. Physical Laws are also general statements; and they are usually held to be capable of

¹The above are paraphrases, unless otherwise indicated; the references are, respectively, as follows: Individuals, pp. 23, 24; 41; 94; 96; 100; 132; 104; 105.

²I am indebted to Professor Bernard Rosen for suggesting the use of this framework or analogy.
being true or false. A physical law, then can plausibly be held to be a true universal generalization which merely describes certain aspects of the physical world. In spite of the fact that physical laws are descriptive, the modal term "necessity" (necessary, must, ...) is often used with lawlike statements. For example, the following are considered appropriate statements of certain physical laws:

1. If $x$ is an unsupported body, then $x$ must fall. OR It is necessary that if $x$ is an unsupported body, $x$ falls.

2. It is necessary that if a species survived in a snowy region, then the members of the species are white.

3. Water, when heated to 100°C. at sea level under normal conditions, must boil or cannot fail to boil. OR It is necessary that water, when heated to 100°C. at sea level, under normal conditions, will boil.

Here, then, we have cases of necessity which seem to be intimately tied to description and which seem to be arrived at not through merely an analysis of meanings (as are analytic necessary truths) but via experience. Says D. Hamlyn writing on the "A Priori and A Posteriori", "Some philosophers maintain that natural laws represent necessary truths, and they do not think this incompatible with the view that natural laws can be arrived at through experience".

He continues,

An even greater number of philosophers would be willing to assert that, in some sense of the word "must", experience can show us that something must be the case... Certainly the "must" in question is not a logical "must", and empiricists have tended to maintain that all necessity is a logical necessity. This, however, is just a

---

Concerning this latter claim there is much controversy. Instrumentalists deny it. The issue is not yet settled.
dogma. It seems plausible to assert that an unsupported body must in normal circumstances fall to the ground.4

In light of the above, can we not see Strawson's use of the modal "necessary" (in his descriptive metaphysical claims) as similarly meaningful and plausible? For example, Strawson claims to be describing, so the claims will be descriptive. Next, Strawson clearly holds them to be capable of being true or false.5 Furthermore, a look at the examples reveals them to be generalizations. And the modal "necessary" occurs in or beside each.

Furthermore, in order to discover and formulate a law of nature, one must both know the language in which the law is formulated and look at the world (see what the nonlinguistic facts are). So, also, in order to discover and formulate a conceptual law (truth), one must both know the language and look at the world (see what the nonlinguistic facts are). For example, concerning a conceptual truth (a descriptive metaphysical statement) involving identification (see d., or e.), one must not only know what it means to identify something but also such facts as, what sorts of things are able to be identified, in what sorts of contexts are they able to be identified, and so on. Or, for example, concerning a., we must know if it is a fact that we are sometimes able to re-identify non-continuous particulars. Concerning b., we must know what death is and what it does and does not involve, in order to discover whether b. is true or not. And so on.


5Individuals, p. xiv.
It thus seems to be that the similarities between physical law statements and their support and what I shall call conceptual laws or truths (accurate descriptive metaphysical statements) are sufficiently striking and numerous to warrant the statement that in so far as the use of necessity with respect to physical laws is intelligible and plausible (which it relatively is), so also Strawson's use of necessity with respect to conceptual laws (or truths) is intelligible and plausible.

Now, returning to our original questions, the first was "What sort of necessity is involved"? The answer is, the necessity involved in physical law statements is unique to physical law statements; it is neither logical necessity, nor psychological necessity, nor practical necessity, but what has been called nomological necessity. Roughly, then, to say that a statement is nomologically necessary is to say that the statement is a true descriptive universal generalization about physical objects (or other scientific objects specifically recognized by scientists).

On the other hand, (to complete the analogy), the necessity involved in conceptual law statements is unique to conceptual law statements; it is neither logical necessity (in the usual sense), nor psychological necessity, nor practical necessity, nor nomological necessity. It is—conceptual necessity. To say that a statement is conceptually necessary is to say that the statement is a true descriptive generalization about concepts or their interrelations or about physical objects, persons, or whatever.
One might wonder what it is that justifies the distinction between nomological necessity and conceptual necessity. I can answer this question while also answering the second of the two questions I originally asked: "How are such necessities discovered?" For, I have distinguished the two necessities because there is a major difference in how each gets discovered or established. While both types do depend for their discovery and formulation upon knowing the language (the conceptual scheme) in which the formulation and discovery occur and observing the world; nonetheless, in the case of nomological necessity, it would seem that the major prerequisite for the discovery of the laws is observation (including experimentation) of the world. On the other hand, in the case of conceptual necessity, it is clear that the major prerequisite for the discovery of the laws is knowing and carefully examining (describing) our linguistic or conceptual scheme. This difference, I wish to suggest, justifies the differences in names; and explains my use of the term "conceptual" for the latter laws. Whether or not there are any further differences between the two, I do not know.

I must mention that my claim concerning nomological necessary statements (above) is not strictly true. Not all true descriptive universal generalizations are accepted as law-like statements. That is, such a statement as "All coins in my pocket are silver" is a true descriptive universal generalization, yet it is not held to be a physical law; rather, it is held to be a mere accidental generalization. One widely accepted way of distinguishing law-like universal generalizations from non-law-like (accidental) universal generalizations
is to inquire if the to-be-judged universal generalization supports its corresponding counter-factual. That is, if the corresponding counter-factual remains true if the original generalization is true, then the generalization is considered law-like; if not, (if putting the original into counter-factual form renders the counter-factual false), then the original is considered an accidental universal generalization.\(^6\)

However, I do not believe that the above effects my case at all, for two reasons.

R.1. The similarities between the actual law-like (non-accidental generalizations) and the conceptual laws or truths are still the same; they both, 1. are descriptive, 2. are capable of being true or false, 3. are universal generalizations, and 4. are supportable via language acquisition and fact-gathering.

R.2. The conceptual claims (laws, truths) are all such that if they are true, their corresponding counter-factuals are also true. For example, a. if the claim "If our conceptual scheme is one of a single spatio-temporal system of material things, then it is necessary that we are sometimes able to re-identify non-continuous particulars on this scheme." is true, then the corresponding counter-factual is true: "If our conceptual scheme were one of a single spatio-temporal system of material things, then it would be necessary that we sometimes be able to re-identify non-continuous particulars in this

scheme." Or take e.g., if it is true that, "If x is a person, then x necessarily has corporeal attributes as well as other kinds of attributes", then it is true that "If x were a person, then x would necessarily have corporeal attributes, as well as other kinds of attributes". And one can go through each of the examples, in the same way. So, since none of the examples fit the analogue of the accidental as opposed to law-like generalization, we need not, with respect to these at least, be concerned.

I submit, then, that the "permanent relationships" which Strawson referred to in the quotation in chapter II are conceptual law relationships such as I have listed. I further submit that the structure which Strawson speaks of as revealed by the descriptive metaphysician is the set of laws or truths (these and others like them) that I have listed.

Lastly, one other aspect of the notion of conceptual law or truth should be mentioned (more about it will occur here and there throughout the dissertation). Strawson is sometimes willing to infer from certain descriptive claims about our conceptual scheme (that is, from certain conceptual truths about concepts) certain descriptive claims about the way the (non-conceptual) world is. For example, as we shall later see, Strawson will infer from the conceptual truth that necessarily the concept of a person is the concept of an entity such that both mental and physical predicates appropriately apply, the further conceptual truth that a person must be a physical entity. (One can, also, of course—though Strawson does not—infer that a
person must be some sort of mental entity as well.) We shall see, however, that Strawson is not always willing to go from talk about concepts (or predicates) to talk about the world. This will turn out to be one of the weaknesses of his theory of self.
CHAPTER IV

THE CARTESIAN THEORY

In this chapter I shall examine the Cartesian theory of self. This theory, and variations upon it, are held and defended even today. Though the Cartesian theory is sufficiently interesting itself, our ultimate aim in examining this theory, as well as others, is to see Strawson's theory as an attempt to produce a theory which is preferable to some major theories of self, including The Cartesian Theory.

In an article entitled, "Self, Mind and Body" Strawson characterizes the Cartesian theory as follows: The Cartesian theory holds that a person is a combination of a Body (a material substance) and a Soul or a Mind (a mental substance). When one speaks about a person, one speaks either about their body or about their mind. That is, there are two types of predicates that are applicable to persons: (1) Body-Predicates and (2) Non-Body (Mental) Predicates. Between

---


2This article appeared in the journal Common Factor, No. 4, (Autumn, 1966), 5-13, seven years after the revised article "Persons" appeared as chapter iii of Strawson's major work, Individuals.
these predicates there is an unbridgeable gulf. No body-predicates properly apply to the mental part of a person. No mental predicates properly apply to the bodily or material part of a person. A person, then, is reducible to (analyzable into) a body and a soul (or mind). More accurately, a person is a soul and a body in some sort of relation. The Cartesian theorist holds furthermore that bodies can exist independently of minds; and vice versa. Lastly, Strawson herein characterized the Cartesian as holding that the concept of a mind (or soul or individual consciousness) is not dependent upon the concept of a person. That is, the Cartesian holds that a concept of a soul or mind is intelligible by itself; it is not necessary to have a concept of a person who has a mind (an individual consciousness) in order to have a concept of a mind.

In order to better grasp this claim let us take an example of a concept which is logically dependent upon some other concept. Let us call the concept which is dependent a non-basic (secondary or derivative) concept; and call the concept upon which this secondary concept depends a basic (or primary) concept. For example, a person is a basic concept in relation to the concept of a person's arm. If one had no concept of a person one could not have a concept of a person's arm. That is, to have a concept of a person's arm it is necessary to have the concept of a person. The Cartesian Theorist, then, is claiming that the concept of a mind or soul is not a derivative concept; it is not a concept that logically could not occur

---

3 Strawson can here be seen as offering a conceptual law or truth. See chapter iii.
without some other (the basic concept) having occurred. Put in still another way, in terms of reference, rather than intelligibility or meaning, the Cartesian Theorist is claiming that in order successfully to refer to a mind (soul, individual consciousness) one need not refer to a person who has that mind.

In chapter iii of *Individuals*, Strawson furthermore describes the Cartesian Theory as claiming that though it may seem that we ascribe body-predicates and mental-predicates to the same entity, a person, this "seeming" is in fact just a "linguistic illusion" (page 89). Says the Cartesian, there is no common owner of these two different types of predicates. When we say things like "I am six feet tall" and "I remember her" it surely seems that the "I" refers to one and the same entity. But, on the Cartesian Theory the first use of "I" refers to a body and the second use refers to a mind. As Strawson puts it in the aforementioned article "Self, Mind and Body", for the Cartesian, "the history of a human being is not the history of one two-sided thing, it is the history of two one-sided things." (p. 6)

The above, then, is Strawson's characterization of the Cartesian Theory. Strawson is careful not to commit himself to claiming an historically accurate account of Descartes' theory. Thus, any extended historical-critical discussion of Strawson's characterization would not be particularly philosophically relevant. However, I would point out that in fact Descartes' philosophy of self is very far from unambiguous. It is true that Descartes sometimes speaks as if a person is

---

4 See the footnote in *Individuals*, p. 89.
indeed a mind or soul and a body, in some sort of relation. But he also often seems to deny this and affirm that a person (an "I") really is merely a mind or a soul. And he most often speaks of a person as a mental substance (not two related substances). Furthermore, two quite recent philosophers who claim to be defending a Cartesian-like theory of self produce theories which hold that a self or person is a mind only; not a mind and a body. For example, H. D. Lewis writes, "... my real self is my mind ...," and "... it is only in a derivative and secondary sense that my body is said to be myself at all." He adds, "My body is not strictly myself, or some part of me. It is something to which I am very specially related, no more." Jerome Shaffer, again claiming to defend a theory "very much like that of Descartes" says that he will argue that "... when I say, 'I weigh one hundred eighty pounds', strictly speaking I am saying something not about myself but about something else--namely, a physical object which is contingently connected with me."

Thus, it is at least arguable that Strawson is dealing only with one of two Cartesian theories of self. However, the one he does deal with is sufficiently important to warrant an examination.

---

6 See, for example, Meditations, pp. 26, 27, 43, 49.
8 Ibid., p. 17.
9 Shaffer, "Persons and Their Bodies", p. 59.
10 Ibid., p. 70.
The theory, then, which I shall examine and treat as "the Cartesian Theory" (hereinafter referred to as CT) will be the Strawson characterization of Descartes' theory. What I wish to do is first, to show that Strawson rejects the CT precisely because it does not, can not, adequately accommodate a part of the phenomena, and because it either cannot solve certain relevant problems or because there are other serious objections to the theory. I will then critically examine Strawson's claims and arguments. And finally, in the interest of completeness, I will judge the theory in relation to those further aspects of Theory Selection (as developed in chapter i) which Strawson omitted.

I will now try, as best I can, to present Strawson's case against the CT. In terms of chapter i and PHENOMENA ACCOMMODATION, Strawson claims that the CT cannot adequately accommodate PH.2., PH.3., and PH.4. That is, Strawson asserts that the CT cannot accommodate re-identification; it cannot explain how we can know that a person now is the same person as the person we met, for example, yesterday. Given this, then, it is obvious that the CT would not be able to account for how there can be clear, ordinary, and generally successful ways of re-identifying persons (PH.2.). Next, Strawson argues that the CT cannot accommodate identification or individuation of persons. Thus, if Strawson is correct, the CT would not be able to account for the phenomena that there are clear, ordinary, and generally successful ways of individuating and identifying persons which we use every day, (PH.4.) and (PH.3).
In my opinion, Strawson is correct. The CT cannot accommodate re-identification, identification, or individuation of persons. However, I will not defend this immediately. My arguments will occur in the upcoming Critical Examination of two of Strawson's major objections (0.1.) and (0.2.) of the CT. The reason for this is that Strawson uses the (alleged) fact that the CT is unable to accommodate re-identification, identification, and individuation as premises in his arguments against (OBJECTIONS TO) the CT. Thus, in order critically to appraise Strawson's objections (specifically to determine the truth of the premises), I will have to determine if the CT is, in fact, unable to accommodate these phenomena.

In terms of PROBLEM-SOLVING or OBJECTION-PRODUCING, Strawson offers two major objections to the CT. Before I present these objections I wish to make it clear that Strawson's arguments as he presents them are far from clear. Thus, I have taken the liberty of expanding Strawson's arguments in an attempt to make them clear and to strengthen them. With the above "in mind", let me now construct what I consider to be Strawson's two major objections to the CT. I will, in the case of both objections, lay the argument out as clearly and succinctly as I can. Then, immediately following this I will explicate the defense, if any, which Strawson offers for each premise.

0.1.
1. If the CT is correct, then one cannot identify other persons.
2. If one cannot identify other persons, then one cannot
ascribe states of consciousness to other persons.

3. If one cannot ascribe states of consciousness to other persons, then one cannot ascribe states of consciousness to oneself.

4. But, clearly, one can ascribe states of consciousness to oneself.

Therefore, the CT is not correct. 11

The Defense:

Premise 1: Strawson is clearly committed to the view that for an entity to be identifiable the entity must in some way be empirically accessible. 12 Now, since the Cartesian person is made up of two entities, two identifications must take place. There must be identification of the body (as, for example, John's body). And there must be identification of the mind (John's mind). Strawson holds that there is no problem with the first (bodily) identification. We have ordinary and generally successful ways of identifying bodily entities. But the second identification is troublesome. We simply to not have any ordinary empirical ways of identifying minds. Certainly, there are no ordinary empirical criteria for identifying minds alone (subjects

---

11I have reconstructed this argument from material on pages 94-97 of Individuals.

12This is best brought out in Strawson's discussion of Kant's critique of Cartesianism found in The Bounds of Sense, (London, Methuen and Co., 1966), pp. 162-169. (Hereinafter referred to as The Bounds of Sense)
of experience, alone). Thus, since one necessary part of the identification of a person cannot take place (namely, mind-identification), other persons cannot (on the CT) be identified.

Premise 2: Strawson offers no defense for Premise 2. He simply affirms, "... surely there can be a question of ascribing only if there is or could be a question of identifying that to which the ascription is made ..."14 On the next page (96), he again affirms without defense that, "One can ascribe them (states of consciousness) to others only if one can identify other subjects of experience."

Premise 3: Strawson elaborates on Premise 3 by saying 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, ..."15 He continues, "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."16

---

13"...one cannot identify others if one can identify them only as subjects of experience, possessors of consciousness." Individuals, p. 96.

14*Individuals, p. 95.

15*Individuals, p. 94. N.

16*Ibid., p. 95. N.
Premise 4: This premise surely is obvious, thus no defense is needed. However, Strawson himself finds it at least puzzling, and thus offers some further explanation. He says that someone might find it puzzling to speak of identifying oneself. In order to ascribe something to something, one must be able to identify that something to which the ascription is made. But because the identification of oneself may seem puzzling, the self-ascription may seem puzzling. Strawson thus feels obliged to show that self-identification does make sense, and thus does not jeopardize Premise 4. Says Strawson, identification of oneself does, after all, make perfectly good sense because, ". . . we speak primarily to others, for the information of others." He continues, "In one sense, indeed, there is no question of my having to tell who it is who is in pain, when I am. In another sense, however, I may have to tell who it is, i.e., to let others know who it is."¹⁷ Thus, there is no puzzle. And, clearly, one can (and often does) ascribe states of consciousness to oneself.

Critical Examination of 0.1.

One way of falsifying Premise 1 would be to show that a Cartesian theorist is able to identify other minds and, thus, other persons. One way of identifying other minds would be to identify them on the basis of the identification of bodies (since there is no problem of

¹⁷ Individuals, p. 96.
identification here). For example, one could identify first one's own mind as "the subject of my experiences which stand in a unique causal relation to my body M." Then identification of another's mind could be "the subject of those experiences which stand in the same unique causal relation to body N as my experiences stand in to body M." Strawson has a response to this attempt. Namely, the identification of one's own mind was to be established by its relation with a body, but the specification of this relation depends upon having already made the identification of oneself ("the subject of my experiences which stand in a unique causal relation to my body"). In other words, in order to identify the subject of experience (the mind) one must already have identified the subject of experience (the person, i.e., the body and the mind). This, then, will not do.

But why could it not be that one has direct knowledge at least of one's own mind and its unique causal relation to one's body M? Thus, one could directly identify one's own mind. And one could identify another's 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." But this also will not do. For, there is no justification for saying, "the subject ..." That is, even in the case of oneself there is no way of knowing if a subject of experience is involved, or two, or three, and so on.\(^{20}\)

\(^{18}\) Individuals, p. 96.

\(^{19}\) Individuals, p. 97.

\(^{20}\) Individuals, p. 97. Also, see the forthcoming Critical Examination of 0.2., especially in relation to its Premise 1.
As Strawson puts it,

We can say to ourselves things like 'I am aware of myself now' or 'This is how it is with me now' and say such things with the conviction of their expressing absolutely indubitable fact. And then perhaps we may begin to feel that we don't have to explain the notions of identity and difference as applied to the soul, for we have direct experience of the individuality and identity of the soul, experience which might be expressed in remarks like these.

But, he continues,

All the immediacy and indubitability of experience which seem to go with the use of 'I' and 'me' in such remarks as I've just quoted could be preserved while re-expressing the remarks in some such form as 'This is a conscious experience' or 'The soul having this experience is conscious of itself as having this experience'. Then it would be apparent where the limits of immediacy and indubitability fall; it would be apparent that there is nothing in the experience itself to rule out the suggestion that there might be a thousand exactly similar experiences occurring in association with the same body--hence a thousand souls simultaneously associated with the body--and equally nothing to rule out the suggestion that the, or each, soul having such temporal series as Kant spoke of--hence, perhaps, a thousand souls the next moment. 21

But, again, perhaps the Cartesian could argue that, for example, one of my experiences is "willing my arm to move and its moving", and if there were other minds standing in this same relation to this (my) body, we would expect many actions resulting from those other minds which would puzzle us about their origin. Yet this does not occur. Therefore, we are justified in holding that there is but one mind associated with our body.

21 Strawson, "Self, Mind and Body", pp. 11 and 12.
As a counter to this argument, however, I would suggest that the argument begs the question. For, reference is made to "willing my arm to move and its moving", but such a claim is precisely what is in question. That is, the argument depends upon a reference to a claim which claim presupposes the identification of oneself—which identification is what is in question. For we are now trying to decide if even in the case of oneself (my), one can know whether identification is possible.

Thus, so far as I can see, CT cannot identify other persons, for it cannot identify other minds, and it cannot even identify one's own mind. Premise 1. is secure. (And the CT is unable to accommodate PH.3.)

What about Premise 2? This premise is not at all obvious, as it stands. For, if "cannot" means, as I believe it does, "as a matter of fact, cannot" then I see nothing at all from preventing someone from accidentally ascribing a state of consciousness to someone else (i.e., without being able to identify him). On the other hand, in relation to the CT, it is surely correct that if other-ascription is often done and often done successfully (as it presumably is), then it would be quite implausible to suppose that this successful ascription occurred again and again accidentally. Thus, generally speaking, it is true that if one could not identify other persons, then one could not have a conception of the appropriate occasions for other-ascription.

Premise 3 can also be worded, "If one can ascribe states of consciousness to oneself, then one can ascribe states of consciousness to others." (This is, of course, the contrapositive of the original
premise.) Strawson is here simply asserting that if, as a matter of fact, one can ascribe states of consciousness to one member of a class (the one member being oneself, the class being the class of persons) and if one has a conception of the appropriate occasions on which to ascribe states of consciousness to this one member, then one also has a conception of the appropriate occasions upon which to ascribe states of consciousness to other members of this class, even though one might never actually ascribe states of consciousness to these other members. This minimum claim, is, I think, not open to objection.

Premise 4 is clearly true.

The argument, then, is not only valid but is sound. Thus, Strawson has produced at least one good objection to the CT.

0.2.

1. If the CT is correct, one cannot individuate or re-identify minds.

2. But if one is able to speak coherently about something (minds), then one (must be) is able to individuate and re-identify minds.

3. Clearly, one is able to speak coherently about minds (or individual consciousness).

Therefore, the CT is not correct. 22

22I have reconstructed this argument from material on pages 9 through 12 of Strawson's article, "Self, Mind and Body."
The Defense:

Premise 1: Strawson's defense for Premise 1 is largely a challenge thrown to the Cartesian. Strawson simply does not understand how a Cartesian can individuate and re-identify minds, without abandoning his position. He writes, for example, "If I were to suggest that when the man, Professor X, speaks, there are a thousand souls simultaneously thinking the thoughts his words express, . . . how would he persuade me that there was only one such soul?" Further, "(How would each indignant soul, once the doubt has entered, persuade itself of its uniqueness?)" The major problem, once again, is that while there are ordinary empirical ways of individuating and re-identifying (after a lapse of time) physical entities, there are no ordinary empirical ways of individuating and re-identifying non-physical entities (minds).

Premise 2: This premise is asserted without defense. It is apparently to be taken as an obvious premise.2\textsuperscript{23}

Premise 3: This is another such premise.

Critical Examination of 0.2.

Is it the case that the Cartesian Theorist is unable to individuate and re-identify minds? Why, for example, couldn't the Cartesian use the granted ability to individuate and re-identify bodies as a means, indirectly, of individuating and re-identifying minds? For

\[23\text{Strawson, "Self, Mind and Body", p. 10.}\]

\[24\text{Strawson, "Self, Mind and Body", p. 9.}\]
example, why not hold that the following is true, "one body, one mind"; "same body, same mind". Since one can know if one rather than two or three, (and so on) bodies are involved, one can thereby know if one rather than two or three (and so on) minds is involved. In short, could one not provide for individuation and re-identification of minds via individuation and re-identification of bodies? The answer is NO; the Cartesian could not. Because, there is no reason at all why the Cartesian is justified in claiming, "one body, one mind"; "same body, same mind". In order to justify these claims he would have to already have been able to individuate and re-identify minds (for the claims are "one body, one mind"; "same body, same mind")!

Perhaps the Cartesian can individuate and re-identify minds by individuating and re-identifying persons. For example, one might hold that once one knows that one person, John, is here, one thereby knows that one mind, John's mind, is here. And likewise for re-identification. In short, one could hold that "one person, one mind" is true; and "same person, same mind" is true. But, again, the Cartesian cannot solve his problem in this way. For, in order to individuate and re-identify a Cartesian person (body and mind) one must first individuate and re-identify both the body and the mind! In short, on this second attempt, one would have to first individuate and re-identify mind x in order to individuate and re-identify mind x!

But perhaps one has a direct experience of the individuality of one's mind and of its sameness? But even this will not help the
Cartesian. For, we clearly do not have direct experience of the individuality of the minds of others. And, secondly, I argued previously that a person does not have direct experience of the individuality of even his own mind. Thus, Premise 1 is true (and the CT cannot accommodate PH.1. or PH.2.).

Concerning Premises 2. and 3., I have only one reservation; namely, the term "coherent". If by this term Strawson means "meaningful" then I think both premises are true. For, it does, at least, seem clear that one cannot speak meaningfully about some particular object if one is unable to tell even that it is a particular (one) object. Indeed, I would suggest further that if one cannot individuate the object in question (if one cannot even tell that it is an object), then one simply does not even have a concept (an idea) of the object, let alone be unable to speak meaningfully about it.

O.2. then, at least as I have constructed it, is a sound argument and is thus a second good objection to the CT. It should be added that the failure of the CT to be able to accommodate re-identification is also its failure to be able to solve PR.5. For, if the CT cannot even make intelligible how persons can be re-identified, it surely cannot explain in what sense a given person is the same person whom we observed before.

I wish now to examine the CT in the light of the full Theory Selection structure exposited in chapter 1. I have argued that the CT cannot adequately accommodate PH.2., PH.3., and PH.4. Can the CT accommodate PH.1. and PH.5.?
Concerning PH.1, the question arises, can the CT accommodate the agent-character of personhood? Can a Cartesian person do things and have things done to him? With these phenomena, the Cartesian is in a somewhat peculiar position in the cases of a Cartesian person doing something to other persons, and having things done to him by other persons. Namely, though a Cartesian self may be able to affect and be affected by others, he can never know (granting the last paragraph) whether or not he has been affected by or has an affect on, for example, John. For, he cannot know if it is John. Furthermore, he can never know if he has done anything; for he cannot know that it is he who has done it.

In short, then, while a Cartesian person might be able to do things and have things done to him, neither he nor we can ever know if he can. So, we cannot say if CT can or cannot accommodate PH.1. We cannot know. But this fact itself puts CT in a position inferior to that of a theory which clearly can accommodate PH.1. Thus, we shall treat CT (in relation to a theory which can accommodate PH.1.) as failing with respect to PH.1.

Moving on to the Ordinary-Language phenomenon, (PH.5.) one of the most interesting aspects of the Cartesian theory is its ability to accommodate much person-talk. For example, one of the most prominent aspects of person-talk is that persons are most often distinguished from things. As Arthur Danto writes in the Encyclopedia of Philosophy (1967), "That a person is distinct from a (mere) thing, and that any human being, insofar as he is a person, is in consequence
of this status to be treated in a special manner, are two of the main logical features of this concept, . . . "25 The Cartesian theorist has a way of explaining why the logic of person-talk is different from the logic of thing-talk; it is because persons are unique in having minds (or souls). Take, for example, utterances such, "God is often considered to be, in some sense, a person". The Cartesian can account for statements of this sort by pointing out that what is often meant by such statements is that God, too, is considered to be something more than a thing; he is considered also to have a mind (to have an intellect). An utterance such as "John is a person; not just a thing" reflects the pro-evaluative element to the logic of much person-talk and again the Cartesian can hold that this evaluative element arises from the fact that persons have this unique entity--mind or soul.

On the other hand, as a result of its failure to accommodate PH.2., PH.3., and PH.4., there are a number of ordinary true statements which CT cannot accommodate. For example, "John is thinking" could not be known to be true on CT, for we cannot know if it is John who is thinking, or Jack, or Jill, and so on. Another example would be the failure of CT with respect to the statement, "There are two persons over there." Since CT cannot individuate persons, it cannot grant that this claim is obviously true. A still further example of CT's failure could be instanced by the statement "That is Joan." For,

the CT is unable to re-identify persons. Thus, it could not be obviously true at all that, "That is Joan."

Thus, CT cannot satisfactorily accommodate PH.5.

Let us now look at the CT in terms of PROBLEM-SOLVING. I have suggested that the CT cannot solve PR.5. But what about PR.1, PR.2, PR.3, and PR.4?

PR.1.: The CT does indeed specify carefully the nature of a person. As has been said, a person is a two-entity "entity". The two entities are mind and body and they are radically distinct. One entity has physical characteristics; the other does not, and so on.

PR.2.: The CT does indeed have a problem with PR.2. The so-called "problem of other minds" is precisely the problem of explaining how we know that there are other minds than our own and thus how we know that there are other persons (bodies with minds). The problem that the Cartesian has is that the minds of other persons are not empirically accessible; while one may be able to know that oneself has a mind, by direct knowledge (direct awareness), one has no such direct access to the minds of others. How then can one know others exist? The Cartesian, it seems, must claim that some sort of indirect knowledge of others is possible. Now the major attempts to defend some such approach have been made via some form of analogical argument. Whether or not some form of analogical argument will work

26If my earlier arguments are correct, however, one does not even have direct awareness of one's own mind. Thus, an analogical argument approach cannot even begin to be generated.
is one of the most controversial issues in contemporary philosophy of mind. In short, it simply cannot be said that a generally accepted solution via analogical argument has been produced. However, perhaps there is some way, other than via analogical argument, that the Cartesian can solve this problem. For example, a somewhat modified Cartesian theory might be said to escape this problem. One could hold that minds are "posits" (entities said to exist not because they are directly known or "observed", but said to exist because their existence would account for a certain range of phenomena—much like scientific "posits" such as atoms, posited to account for a range of physical phenomena). Now, the Cartesian theorist (of this sort) could argue that the mind is such a posit; and the range of phenomena which it accounts for is a mass of behavioral phenomena. Examples of this phenomena could be: the ability of persons to play chess, solve complex math problems, and generally to respond in various ways that other entities (non-persons) cannot respond. Then, the modified Cartesian can solve the problem of other minds and thus the problem of other persons. For, he now can respond that it is on the basis of certain types of behavior (which are publicly accessible) that one can know that one's own mind and the minds of other persons exist.

I see, however, at least two major objections to this attempt to retain a kind of Cartesian theory of self and yet solve PR.2.

0.1. One would not need to posit minds in order to adequately accommodate the relevant behavioral phenomena. A brain would equally well satisfactorily account for these phenomena. And the brain has the major advantage of being publicly observable. Furthermore, there is empirical evidence that this entity (brain) is closely connected to the aforementioned behavioral phenomena.

0.2. Furthermore, there is a growing body of literature making a case for the claim that there is no class of behavioral phenomena which non-persons cannot do. That is, it is controversial whether there even is the unique-to-man phenomena which the modified Cartesian alleges. 28

PR.3. In discussing PR.3. and throughout this study, I will use the term Solipsism to refer (as it usually does) to the view that there exists only one person; namely oneself. That would mean that if one could show that there exists any entity (including a person or persons) other than oneself, Solipsism would, thereby, be refuted. If, then, one could solve the problem of other minds (that is, if one could show how we know that there do exist other minds and thus other persons) one would thereby have refuted Solipsism (that is, shown it to be false). On the other hand, one could solve the problem

28 It should be mentioned, however, that the plausibility of this claim is being widely questioned today. See, for example, the growing literature on the abilities of robots and computers. For example, see articles such as Michael Scrivens, "The Compleat Robot: A Prolegomena To Androidology", reprinted in Dimensions of Mind, edited by Sidney Hook as well as articles by, for example, A. M. Turing and Hilary Putnam.
of Solipsism, and yet not have solved the problem of other minds. For, it might be that one could show that entities other than minds or persons exist, thereby refuting Solipsism, but be unable to show that other minds exist, thereby failing to solve the problem of other minds. Given this, in so far as the CT has been shown to be unable to solve the problem of other minds, it does not follow that the CT is thereby necessarily unable to solve the problem of Solipsism. In short, from the fact that the CT generates the problem of other minds (and thus of other persons), it does not follow that it generates as well the problem of Solipsism.

Granting the above, the question still arises, Can the CT solve the problem of Solipsism? Again, granting that the Cartesian Theorist is unable to solve the problem of other minds, his only hope is to show that a Cartesian self can know that an entity (at least one) other than a mind or a person does exist. At this point, I think the Cartesian is, again, in trouble. For, first of all, the predicate "knowing" would surely be a mental predicate and thus is properly predicable only of the "I" when the "I" refers to the mind or soul. Next, if the mind or soul is, as the Cartesian claims, non-physical and if the "world" which is known is, as it presumably is, physical, then the Cartesian has a major epistemological problem, quite analogous to the mind-body problem. Namely, the question arises, how can there be any sort of intelligible interaction between a non-physical entity and a physical entity sufficient to generate a "knowing"; indeed, how can there be any intelligible effective relation between the two?
In sum, then, the Cartesian cannot solve PR.3. Nor does the Cartesian argue that these phenomena (these "knowings") are illusory.

PR.4. The differences between the mental and the physical aspects of persons are always carefully drawn by the Cartesian theorist. The physical aspect (the body) of a person is measurable, publicly accessible, in short (to use a term used by Descartes himself) extended; the mental aspect (the soul or mind) is non-measurable, private, in short, non-extended. As a contemporary defender of Descartes puts it, "My mind has neither height nor length nor breadth . . . tallness and so one has nothing to do with my mind, but only with my body. The much maligned Descartes was obviously right in maintaining that it was distinctive of minds not to be extended." 29

On the other hand, whether a Cartesian theory of self is able to solve the mind-body relationship problem is still very controversial. However, it can be said (and even this is damaging to the CT) that no generally accepted solution has been offered to date. Furthermore, in recent years, a rival theory, the Identity Theory (there are various versions by now) has become a leading competitor as a solution to the age-old mind-body problem.

Next and lastly, let us examine the CT in terms of III. (GENERAL THEORY PREFERENCE PRINCIPLES). First (P.1.), it is clear that

in so far as the CT characterizes a person as involving two entities, a simpler theory is possible. We shall see later that Strawson claims to have produced just such a simpler theory.

Concerning P.2., one of the less satisfactory aspects of the problem of self is that there are no well-established theories, scientific or otherwise, which relate in any direct way to this problem. One cannot, for example, find any general consensus of expert opinion in psychology concerning the self. One finds, rather, a mass of "schools" and "movements" (Freudian, Adlerian, Jungian, Existentialistic, Behavioristic, and so on) with a corresponding plethora of varying views on the self and varying technical uses for such terms as "self", "ego", "person", and so on. Thus, P.2. must, at this time, remain idle.

Lastly (P.3.), historically, at least, it is clear that Cartesian theorists who attempted to solve the problem of mind and body (which is part, you recall, of specifying the nature of a self) did "solve" this problem only by using ad hoc devices. For example, one calls to mind Descartes infamous pineal-gland "solution". As well, those who later preserved the self as mind and body but "solved" the problem of mind and body by, for example, the doctrine of occasionalism, again simply used an ad hoc device. Even today, no mind-body dualist has produced a generally accepted explanation of the relation between mind and body. In so far, then, as the only "solutions" available are ad hoc, the Cartesian theorist is burdened with still another "mark" against his theory.
Enough has been said, I think, to reveal that the CT is very unsatisfactory. It has failed to satisfactorily accommodate any of the phenomena. It has failed to solve problems PR.2. through PR.5. As well, it has two further unresolved objections against it; and, thus, has two additional unsolved problems. And, it has failed in simplicity (P.1.), at least, in so far as a simpler theory is possible. And it failed in the area of ad hoc additions (P.3.).

On the other hand, CT might still be the best theory now available. In order to decide this, we must look at some other major theories of self. I shall next examine Hume's famous "bundle" theory of self.
HUME'S THEORY

In this chapter, I shall be concerned with Hume's famous "bundle-theory" of self. The structure of the chapter will be as follows: first, I will carefully explicate Hume's theory; secondly, I will present and critically respond to Strawson's critical remarks on Hume's theory; thirdly, I will examine the success of HT in relation to PHENOMENA ACCOMMODATION, PROBLEMS, and the EXTRINSIC PREFERENCE PRINCIPLES. And lastly, I will offer a short comparison of the two theories so far examined (CT and HT).

According to Hume, a self (a person, a mind) is a set of perceptions in relation. Says Hume,

For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception. When my perceptions are remov'd for any time, as by sound sleep; so long am I insensible of myself, and may truly be said not to exist. And were all my perceptions remov'd by death, and cou'd I neither think, nor feel, nor see, nor love, nor hate after the dissolution of my body, I shou'd be entirely annihilated, nor do I conceive what is farther requisite to make me a perfect non-entity. If any one upon serious

1HT will refer to Hume's theory of self.
and unprejudic'd reflections, thinks he has a different notion of himself, I must confess I can reason no longer with him.²

He continues, "The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations."³ Furthermore, Hume asserts that, "... the true idea of the human mind, is to consider it as a system of different perceptions of different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other."⁴ And, lastly, Hume observes that, "When I turn my reflexion on myself, I never can perceive this self without some one or more perceptions; nor can I ever perceive any thing but the perceptions. 'Tis the composition of these, therefore, which forms the self."⁵

Negatively, Hume rejects any theory of self which considers the self as a substance (a permanent entity) which has perceptions but which is different from the perceptions. The term "perception" is used by Hume as a generic term. Specifically, perceptions can fall into one of these four categories:

1. impressions of sensation (such as "seeing a brown chair" or "hearing sound z")

---


³*Treatise*, p. 253.

⁴*Treatise*, p. 261.

⁵*Treatise*, p. 634.
2. impressions of reflexion (such as "desiring x" or "hoping y")
3. ideas of sensation (such as "one's idea of the brown chair" or "one's idea of the sound z")
4. ideas of reflexion (such as "one's idea of desiring x" or "one's idea of hoping y")

For Hume, ideas are simply "faint copies" of impressions. That is, the sole distinguishing mark between ideas and impressions is that impressions are more "forceful", more "vivacious", than ideas.

Lastly, these perceptions that constitute a self are related in various ways. The relations involved are the following: resemblance, contiguity, and causation. That, then, is Hume's theory of self. To sum up, a self is a set of perceptions (impressions and ideas, of sensation and of reflexion) related in various ways (by resemblance, contiguity, and causation).

It should be pointed out that, just as there are contemporary defenders of CT or variations upon it, there also are contemporary defenders of HT or variations upon it. For example, Bertrand Russell in *The Problems of Philosophy* agrees with Hume that a self is not a permanent entity of any sort, and, further, like Hume, suggests that, "When we try to look into ourselves we always seem to come upon some particular thought or feeling, and not upon the "I" which has the

---

thought or feeling." Nonetheless, Russell does not, in this early work (1912) commit himself fully to a Humian theory. However, by 1936, Russell is very close to Hume. Says Russell, in an essay entitled "Do We Survive Death?", "All that constitutes a person is a series of experiences connected by memory and by certain similarities of the sort we call habit."

Still another example of the contemporary survival of this type of theory of self is the position of A. J. Ayer. For example, in Language, Truth and Logic Ayer denies that a self is a substance and claims to "... have vindicated Hume's contention that it is necessary to give a phenomenalist account of the nature of a self...

More specifically, Ayer holds that a self "must be held to be a logical construction out of sense-experiences." The details of this position I, of course, can not go into. The major point I am making is that in so far as we are examining Hume's theory of self, we are not examining a theory of merely historical interest. Humian theories of self are still "living" competitors for the correct (the best) theory of self.

---


12Ibid., p. 125.
Surprisingly, and unfortunately, Strawson's critical comments on HT are neither extensive nor carefully and systematically developed (especially as compared to his critical remarks on CT). But he clearly does reject HT as inadequate.\(^{13}\) First, Strawson suggests that Hume failed to "find" his "self" (when Hume was introspecting) because the entity he was searching for was precisely "the illusory primary concept of the pure consciousness, the ego-substance . . . of Descartes."\(^{14}\) But more importantly, Strawson is very critical of Hume's search for a principle of unity, after Hume adopts the theory that a self is a set of related perceptions. Writes Strawson,

the problem that does not exist is the problem that seems to have perplexed Hume; the problem of the principle of unity, of identity, of the particular consciousness of the particular subject of 'perceptions' (experiences) considered as a primary particular. There is no such problem and no such principle. If there were such a principle, then each of us would have to apply it in order to decide whether any contemporary experience of his was his own or someone else's; and there is no sense in this suggestion.\(^{15}\)

Strawson further expands this objection in The Bounds of Sense. There he states,

When a man (a subject of experience) ascribes a current or directly remembered state of consciousness to himself, no use whatever of any criteria of personal identity is required to justify his use of the pronoun 'I' to refer to the subject of that experience. It would make no sense to think or say: This inner experience is occurring,

\(^{13}\)See, for example, Individuals, pp. 98-99, 132-133; and The Bounds of Sense, pp. 165, 169-170.

\(^{14}\)Individuals, p. 98.

\(^{15}\)Individuals, p. 132.
but is it occurring to me? (This feeling is anger; but is it I who am feeling it?) Again, it would make no sense to think or say: I distinctly remember that inner experience occurring, but did it occur to me? (I remember that terrible feeling of loss: but was it I who felt it?) There is nothing that one can thus encounter or recall in the field of inner experience such that there can be any question of one's applying criteria of subject-identity to determine whether the encountered or recalled experience belongs to oneself--or to someone else.¹⁶

It would seem to me that Strawson's remarks quite miss the point. Hume's concern was not with the problem of self-identification. Hume could and, I think, would agree with Strawson, as I do, that one does not need any principle and its application to determine whether any given perception is our perception. Furthermore, while I think Strawson is correct that Hume's "looking inward" for his self was a "looking-for" and "not-finding" a Cartesian ego-substance, Strawson is not correct in asserting that no problem of unity exists for Hume. What precisely, then, is Hume's problem (of unity)? It seems clear to me that Hume was bothered by the conflict between (1) his bundle-theory which holds that a self is a set of distinct entities (any one of which could exist separately, i.e., by itself), and (2) the clear phenomenon (recognized by Hume) that a self is an entity. Says Hume, "... having loosen'd all our particular perceptions, when I proceed to explain the principle of connexion, which binds them together, and makes us attribute to them a real simplicity and identity; I am sensible that my account is very defective, ..."¹⁷ Hume's problem is, How is it that one

¹⁶The Bounds of Sense, p. 165.
¹⁷Treatise, p. 635.
can explain the unity and sameness (individuality) of a self (which Hume accepts as a datum) if one analyses a self as a set of distinct perceptions related only by what might be called psychological relations (as opposed to what might be called real relations, or as Hume puts it, real connexions). This is Hume's problem of unity and it makes sense, it seems to me, of the following crucial passage in Hume:

In short, there are two principles, which I cannot render consistent; nor is it in my power to renounce either of them, viz. that all our distinct perceptions are distinct existences, and, that the mind never perceives any real connexion among distinct existences. Did our perceptions either inher in something simple and individual, or did the mind perceive some real connexion among them, there would be no difficulty in the case.  

Now the contradiction that Hume is bothered by is not a contradiction between the two italicized statements--there is no contradiction between those!--rather Hume is, admittedly misleadingly, referring to the contradiction between (1) holding that a self has a unity and an individuality, yet (2) holding also that a self is a set of distinct (separate; not united) perceptions which are not related by any real connexion (which could provide the unity necessary). This is a plausible interpretation for at least two reasons: First, it saves us from having to assert what seems very implausible, that Hume was too stupid to see that the two italicized statements were not contradictions. Secondly, the interpretation I have adopted makes sense of the context. Just prior to these italicized statements is Hume's admission that there is a gap between his perception—

18Treatise, p. 636.
bundle-theory and our attributing to the self "a real simplicity and identity." And immediately following the italicized statements, Hume says that there "would be no difficulty in the case" if we could either show our perceptions did inhere in something simple and individual (in which case there would be no contradiction because Hume could continue to affirm the distinctness of perceptions yet also take care of the unity and individuality of the self by referring to the substance in which they inhere); or, show that there really was some real connexion between the perceptions (in which case there would be no contradiction because Hume could continue to affirm the distinctness of perceptions yet also take care of the unity and individuality of the self by referring to the real connexion(s) uniting the separate perceptions).

That, then, is Hume's problem of unity, and contrary to what Strawson asserts, it is a very real problem for HT. Indeed, Hume hints, as a result of this real problem, that there is something wrong with (defective in) his theory of self. How very correct he was, I shall now begin to establish.

I wish now to examine HT (as I did CT) in relation to the structure developed and defended in chapter i. First, then, how does HT fare with respect to I. (PHENOMENA ACCOMMODATION)? Concerning PH.1., it would seem that Hume could give a relatively plausible phenomenalistic account of "doing". Take an example of raising one's arm. Hume presumably would analyze this into a set of perceptions;

---

19Treatise, p. 635.
impressions of sensation, the arm (a physical object, which is analysed into sets of perceptions) and the idea (a perception) which caused (in the Humeian sense of the constant conjunction of perceptions, of course) the raising of the arm. And I would suppose that the motion involved in the doing (raising the arm), would be analysed merely in terms of the alteration of sets of similar perceptions. That is, the motion of raising one's arm is phenomenalistically analysable on the model of a motion picture film. Each second (split second, . . .?) one stops the film (of a person and his arm), and this is analysable into a set of perceptions (in a certain spatial location—if that makes sense). Next instant (the arm position has minutely altered), another set of perceptions is involved. And so on.

As I said, Hume can thus provide a relatively plausible account of doing. I stress relatively because I am very bothered with the notions of space and time with respect to this phenomenalistic account. But the notions of space and time have not been adequately analysed or explained by any theory (phenomenalistic or otherwise). Thus, in this latter context, Hume can provide a relatively plausible account of doing and thus accommodate PH.1.

The failure of HT to accommodate PH.2. is admitted even by Hume.20 What, though, of identification (PH.3.)? Is there any way of identifying a self if Hume's theory of self is correct? This question properly breaks down into two more. First, can one identify

20Treatise, pp. 635, 636.
oneself? Secondly, can one identify others? Concerning the first, it would seem that there is no problem. For, one has direct experience of oneself every second that one is conscious. For, every conscious experience is a part of oneself. There can, then, never even be a question of not being able to identify oneself, if by that is meant, not being able to decide whether any experience had is an experience had by me (by oneself). On the other hand, it seems clear that a necessary condition of identifying any given thing x is that one be able to distinguish that thing from other things. In other words, there must be some properties (or at least a property) different between x and y, for one to be able to identify x as x or y as y. Now the point is, in HT there is nothing which distinguishes oneself from anything else. Every experience which one has is a part of oneself, including the experience of tables, chairs, other persons, and so on. Thus, identification is ruled out if HT is correct. Thus, HT cannot accommodate PH.3. There is still another problem with identification of others, on HT. Namely, one surely cannot identify x if x cannot even be known to exist. And I shall soon argue that HT cannot solve PR.2. and thus, for this reason too, HT cannot accommodate PH.3.

Next, concerning Individuation, (PH.4.), I have three arguments.

0.1. If one cannot even tell that there are other persons (cannot solve PR.2.) then one surely cannot tell how many other persons there are. I shall soon argue that HT cannot solve PR.2.

0.2. As well, if there is nothing which enables one to know that a person x is the same person x, then, again, how can one know
that there are, for example, six persons, here, now. That is, what seems to be a person x, might be 100 persons or 1000; unless one has some means of distinguishing between persons, or some means of saying x and y are not the same person, then one simply cannot tell how many persons there are.

0.3. Still further in so far as Hume can provide no principle of unity, he has as well failed to provide any means of telling if a self is a self, or two, or more.

Lastly, (PH.5.), how does HT fare with respect to the vast set of meaningful and true claims we make about persons in everyday life (and, of course, in ordinary language). Once again, it would seem that in one way or another, a relatively plausible phenomenalistic account could be given. For example, take the claim that "John is now thinking about it". This is analysable into a set of perceptions and relations, one sub-set of which perceptions is the set of thoughts John is "having". Or take the claim that, "John is going to the show with a group of other persons". John can be phenomenalistically analysable. And, even if solipsism is not able to be escaped by HT, there still is a peculiar sense in which Hume can refer to a group of persons (as one's own perceptions) and phenomenalistically analyse this. The same sorts of phenomenalistic analysis can be done with respect to hoping, dreaming, walking, and so on.

I should point out, however, that in so far as HT has failed to accommodate PH.2., PH.3., and PH.4., there is a sense in which, in spite of its relatively successful phenomenalistic analyses, it still fails to be able to satisfactorily accommodate PH.1. and PH.5.
For, just as on CT one is unable to tell in any given case if a Cartesian person (John) is doing something (since one cannot re-identify, identify, or individuate with respect to persons), the same is true of HT. Likewise, statements such as "John is thinking", "That is Joan", "There are two persons over there", and so on, cannot (with HT as with CT) be granted to be obviously true, for one cannot know if it is John, or Joan, or two persons (since HT cannot accommodate PH.2, PH.3, and PH.4).

I will now examine HT with respect to II. PROBLEMS TO BE SOLVED.

PR.1. HT does carefully explicate the nature of the self. A self is a set of perceptions in relation. Perceptions are either impressions of sensation or reflexion, or ideas of sensation or reflexion. The relations involved are resemblance, contiguity, and causation. It is also clear that Hume rejects positions like: the self is a mental substance; the self is a physical substance; the self is a process; and so on. Furthermore, there is a sense in which, for Hume, a self is not a permanent entity. Rather it is constantly changing; for, its constituents are rapidly perishing existences.

PR.2. Can Hume explain how we know that there are other persons? First, Hume did not recognize the various views possible—consistent with his theory of self—as to the nature of "other persons". For example, from one's own point of view, other persons are a sub-set of one's own private perceptions. On the other hand, from some other person's point of view, oneself (as an other person) is a set of the other person's private perceptions. Is, then, an "other
person" a sub-set of my perceptions? Or is an "other person" a set of his (the other person's) private perceptions? Or perhaps both? Hume does not provide us with an answer. Let us examine each of these views, however. (1) If an other person is a set of my perceptions, then Hume can explain how we can know of the existence of other persons. Namely, we perceive them (that is, a sub-set of my perceptions is them!). But, while this view is open to Hume and does solve the problem of explaining how we know of the existence of other persons, nonetheless, this view is very implausible. For, this view implies that one could literally make other persons go out of existence by oneself not perceiving them. Similarly, only those persons exist which oneself is perceiving. These views are, I take it, clearly false. But furthermore, if this view is correct, it is logically impossible to be a successful spy (to "see" yet not be seen). Yet this is clearly not logically impossible. Indeed, it clearly sometimes happens that oneself is perceived but does not perceive the perceiver. Hume, thus, cannot plausibly hold view (1).

On the other hand (2), if an other person is a set of his (the other person's) perceptions, then it does not seem to be the case that one could know that other persons exist. That is, one's awareness of other persons comes from one's own perceptions. How then could I know about you if you are a set of your private perceptions. One way would be for me to perceive (have?) your perceptions. But it would seem that it is not even logically possible for me to have your perceptions. For, if I did, they would not be yours,
but mine! And, if they were mine, I would not be perceiving you
(that is, that set of perceptions private to you). Lastly, the view
that you are a set of my perceptions and your perceptions partakes
of the objections to both view (1) and (2). Thus, HT cannot solve
PR.2.

PR.3. Hume clearly has major difficulties explaining how we
know that there is something besides oneself (and other persons).
For example, if there exists nothing other than perceptions and
relations, and if in order to know that there exists something other
than oneself, one must perceive it, the problem arises that our per-
ception of this thing (presumably different from oneself), immediately
becomes a part of oneself (one more perception in the set of percep-
tions constituting oneself). For Hume, there simply is no way of
getting beyond oneself. In short, one cannot know anything other
than oneself. Put somewhat differently, with Hume, as with Berkeley,
it is not the case that our perception of x, and x itself are dif-
f erent. Rather, our perception of x is x. And if every perception
of a person is a part of that person, then every x (which is a per-
ception "of x") is a part of that person. This is solipsism. And,
thus, HT fails with respect to PR.3.

PR.4. The mind-body problem is generally considered to be the
problem of explaining the differences and relationship between two
different sorts of entity, mind and body. But HT does not preserve
this distinction. For HT, there are simply perceptions and relations.
Indeed, one of the major virtues of HT is that it does not generate
the mind-body problem. It might, however, be wondered how a common example of alleged mind-body interaction would be explained on HT. Let us take an example of a person being frightened with the result that this person begins to perspire. On HT, there is an impression of reflexion (fear) and then there follows an impression of sensation (the person's perspiring). The relationship between these two occurrences (both the same type of entity; namely, perceptions) is causal (in Hume's sense of causal). Importantly, even if one took the fear-perception to be the mental aspect and the perspiration-perception to be the bodily aspect (which, of course, is peculiar since this later too is a perception and is thus mental), there still is no problem generated. For, Hume's analysis of causality does not prevent any two entities, no matter how similar or different, from plausibly being causally related.

While the above can be granted, nonetheless, HT is not completely satisfactory with respect to PR. For, first, no explanation is given by Hume of why persons make a distinction between mental characteristics and physical characteristics. Secondly, Hume's analysis of the causal relation is very problematic and cannot be said to successfully solve the problem of a relationship between even different sorts of perceptions. So, while HT is more successful (with respect to PR) than is CT (since it does not generate the problem) nonetheless, HT still does not satisfactorily solve PR.
FR.5. HT notoriously cannot explain the unity, the sameness of a self. There are at least two obvious reasons for this failure: (1) Because of the nature of the entities constituting a person; namely, constantly perishing perceptions; (2) Because of the refusal of Hume simply to admit as a phenomenon the substantival nature of persons. The unity, then, could be grounded in the substance; and specifically, in the person as a physical substance accessible to empirical criteria of unity and sameness.

I now wish to produce some further objections to HT.21

0.1. 1. If HT is correct, then a number of statements like the following are not contingent (are logically true).

(a) John is thinking of his mother.
(b) Jack loves Mary.
(c) Jill looks at her coat.
(d) Jim hears his mother calling.

2. But, these statements clearly are contingent (are not logically true).

Therefore, HT is not correct.

Defense: The reason these statements are logically true on HT, is that each of the subjects of the sentence (John, Jack, Jill, and Jim) are analysable into the set of perceptions "of" each of the persons. For example, John is the perception "thinks of his mother" as well as the other perceptions involved at the time. So, in the case of each of these, the predicate is—as it were—"contained in"

21 Each of the PROBLEMS which HT cannot solve is, of course, an objection to HT.
the subject. To put it somewhat differently, the predicate terms of these statements are merely a partial analysis of the subject terms. Put still differently, the predicate (on HT) adds nothing to the concept of the subject. In short, the statements, if HT is correct, are analytic. Yet obviously, these statements are synthetic, are contingent. Therefore, HT is not correct.

0.2. As Roderick Chisholm points out in a recent article on Hume's Theory, Hume, in his introspections in search of his self, found more than a set of perceptions. Hume writes: "For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe anything but the perception." Asks Chisholm, "How can he (Hume) say that he doesn't find himself—if he is correct in saying that he finds himself to be stumbling and, more fully, that he finds himself to be stumbling on certain things and not to be stumbling on certain others?"

Continues Chisholm,

The difficulty is that Hume appeals to certain evidence to show that there are only perceptions, and that when he tells us what this evidence is, he implies

---


23Treatise, p. 252.

not only (i) that there is, as he puts it in his example, heat or cold, light or shade, love or hatred, but also (ii) that there is someone who finds heat or cold, light or shade, love or hatred, and moreover (iii) that the one who finds heat or cold is the same as the one who finds love or hatred and the same as the one who finds light or shade, and finally (iv) that this one does not in fact stumble upon anything but perceptions.\textsuperscript{25}

Hume found not only perceptions, but also he found "that he found those perceptions . . ."\textsuperscript{26} and this clearly implies that there are perceptions on the one hand and something (a self or person) on the other, which finds the perceptions. In short, a self is not identical to a set of (its) perceptions. HT is not correct.

Our next question is, how does HT fare with respect to the GENERAL THEORY PREFERENCE PRINCIPLES?

P.1. It is clear, that Hume has not produced the simplest possible theory; at least with respect to number of entities. For, Hume analyzes the self into two components; (1) perceptions and (2) relations. Thus, a simpler theory of self is possible.

P.2. Since there are no well-established theories relevant to the theory of self, this Principle, for the present, must remain idle.

So far as I know, no ad hoc devices are used by Hume or his defenders.

P.3., then, is not violated.

Lastly, it might be interesting briefly to compare the relative theoretical success of CT and HT. With respect to I. (PHENOMENA

\textsuperscript{25}\textit{Tbid.}

\textsuperscript{26}\textit{Tbid.}, p. 11.
ACCOMMODATION) CT failed to accommodate PH.1., PH.2., PH.3., PH.4., and PH.5.

HT had exactly the same total failure. With respect to II. (PROBLEMS TO BE SOLVED), CT and HT failed to solve PR.2., PR.3., PR.4., and PR.5. They succeeded only with PR.1. On the other hand, HT was somewhat less unsatisfactory with respect to PR.4.

And finally (concerning III. GENERAL EXTRINSIC PRINCIPLES) neither CT nor HT is monistic (both are in their respective ways, dualistic) and thus in terms of simplicity (P.1.) are similarly successful (or unsuccessful). P.2. remains idle with respect to both. And, while CT failed with respect to P.3., HT does not.

A look at the above reveals, I think, the not too surprising fact that Humian Theory is in certain respects, an improvement over the historically prior Cartesian Theory. But the failures of both surely leave room for a more successful theory.
CHAPTER VI

THE NO-OWNERSHIP THEORY

The structure of this chapter will be as follows: first, a careful description of the No-Ownership Theory (hereafter referred to as NT) as it is understood by Strawson; secondly, a presentation and critical examination of Strawson's objections to this theory; thirdly, an examination of NT and its success or failure in relation to PHENOMENA ACCOMMODATION, PROBLEM-SOLVING, and EXTRINSIC PREFERENCE PRINCIPLES; and lastly, a short comparison of NT with CT and HT.

According to Strawson, both Ludwig Wittgenstein and Moritz Schlick held a theory of self which Strawson calls the "No-Ownership" or "No-Subject" theory of self. This theory is characterized by Strawson as follows. According to NT, it is a linguistic illusion that corporeal characteristics are ascribed to the same thing as are ascribed states of consciousness. There is no common owner or common subject such that both corporeal characteristics and states of consciousness (mental characteristics) are properly ascribed to it. Thus, it is a linguistic illusion (and thus is not really the case) that one ascribes states of consciousness to anything at all.
Yet Strawson sometimes describes NT as a dualistic theory; namely, a dualism of one subject, a body; and of one non-subject.\textsuperscript{3} Here, Strawson interprets NT as suggesting that there are two uses of 'I'. One use is when 'I' has corporeal characteristics predicated of it and the 'I' here denotes a body; the second use is when 'I' has a state of consciousness predicated of it and this use denotes a non-subject. Says Strawson, "... both the Cartesian and the no-ownership theorists are profoundly wrong in holding, as each must, that there are two uses of 'I', in one of which it denotes something which it does not denote in the other."\textsuperscript{4} Now presumably to denote a non-subject is simply not to denote (rather than to denote a peculiar sort of entity, a non-subject). Thus, these claims concerning the two uses of 'I' are consistent with the prior description of NT.

Strawson now, however, suggests that there is a sense in which the No-Ownership theorist could admit that states of consciousness are ascribable to something. Namely, they can, in a sense, be ascribed to bodies. Unfortunately, however, such ascription is always "improper".

So far, then, according to Strawson, the No-Ownership theorist apparently asserts that, properly speaking, there is no common owner of bodily and mental characteristics. Properly speaking, no states of consciousness are ascribable to anything at all. On the other
hand, improperly speaking states of consciousness do have an owner; namely, a body. "Infelicitously speaking", states of consciousness "belong to" one's body.

The use of "improperly", "properly" and "infelicitously" here does need some explanation. Apparently the No-Ownership theorist is claiming that since there is no owner or possessor of states of consciousness (Wittgenstein: "the 'I' does not denote a possessor"; Schlick "...the data have no owner or bearer"; quoted in Individuals, footnote, p. 90) the ascription of states of consciousness to something cannot be correct. That is, "properly speaking" seems to be replaceable by "correctly speaking" for, if we were speaking correctly or truly, we would not ascribe states of consciousness to anything; for there is no owner of states of consciousness. But, the No-Ownership theorist is willing to grant that--still improperly speaking--but nonetheless closer to the truth, is the ascription of states of consciousness to a body, rather than to an Ego. That is, the major concern of the No-Ownership theorist is to deny that there is such a thing as a non-bodily entity called an Ego and to deny that there is any sense of "belong to" which is different from "belong to" meaning "is causally dependent upon". The No-Ownership theorist denies, then, that there is any sense of "own" or "belong to" such that it is logically true that "My Ego owns all its states of consciousness". (Let us designate this "belonging to" or "having" as "belonging to \( \_1 \)" or "having \( \_2 \)" and "belonging to" or "having" meaning "causally dependent upon" as "belonging to \( \_1 \)" or "having \( \_1 \)".)
Given that it is the existence of an Ego and the existence of such a relationship as "belong to \( \text{2} \)" (which is different from "belong to \( \text{1} \)" meaning "is causally dependent upon") which the No-Ownership theorist wishes to deny, he is willing to admit that--still improperly speaking--but at least intelligible is the view that states of consciousness "belong to \( \text{1} \)" (in the sense of "are causally dependent upon") a body. For, there do exist bodies (there do not exist Egos); and there do exist causal relationships (there do not exist, claims this theorist "belonging to \( \text{2} \)" relationships). In other words, if we were to speak properly (correctly), states of consciousness are not ascribable (do not "belong to \( \text{1} \)" or "belong to \( \text{2} \)") anything at all. But improperly speaking (speaking incorrectly but intelligibly, since bodies do exist and "belong to" does exist), states of consciousness are ascribable to ("belong to \( \text{1} \)") bodies.\(^5\)

That, then, is NT. In sum, it seems to be the position that a person is a body; nothing more nor less. Bodily predicates are, of course, properly ascribed to persons; for, one is simply predicating a bodily characteristic of a bodily entity. Mental predicates, however, are not properly ascribed to persons. They are not properly ascribed to anything at all. However, they are improperly ascribable to persons (that is, to bodies). Such ascription means in every case merely that "such and such mental characteristic is causally dependent upon such and such a person (body)."

\(^5\) *Individuals*, pp. 90-91.
Strawson is not greatly concerned with whether the theory as developed above is a totally accurate reflection of the views of Wittgenstein and Schlick. Neither shall I be concerned with the issue of historical and interpretive accuracy. Let me merely point out however, that Strawson is surely going beyond the expressed views of Wittgenstein and Schlick when he speaks of "linguistic illusions" and "infelicitous" use of certain words. Nonetheless, it is with the theory as interpreted by Strawson that I shall deal in this chapter.

Strawson has two major objections to NT.

0.1. The illusion of an Ego and of a relationship of "belonging to 2" (the two things the existence of which is denied by the No-Ownership theorist) arises from persons slipping from a claim like (1) "All my experiences are had by (i.e., are uniquely causally dependent upon the state of) body B", which claim is contingent, to the non-contingent claim that (2) "All my experiences are had by E." This latter claim involves both the sense of belongs which the No-Ownership theorist denies and the Ego (E) which the No-Ownership theorist denies. We are not told why this "slip" occurs. But, that it occurs, is what allegedly generates the problem (the belief in E and "belongs to 2").

Now, what Strawson does is try to show that the No-Ownership theorist cannot even speak of the generation of this "slip" or illusion from a contingent claim to the non-contingent claim, without relying upon the very sense of possession or belonging to ("belonging to 2") whose existence he claims to deny.
What Strawson is going to argue is, first, that the original contingent claim, in order for it to remain contingently true, must involve reference to the person whose experience is involved. For, for example, if one eliminated such personal reference from (1) not only would the meaning of the sentence be different, the sentence would no longer be contingently true; it would become contingently false. Why? Because while the original sentence reads (1) "All my experiences are had \( \&) by body B" the revised de-personalized sentence would read "All experiences are had \( \&) by body B" which latter statement is contingently false.

So, the first point of the argument is that the personal reference must remain in statement (1) in order for the illusion to arise from this (1) contingently true statement (which is what the No-Ownership theorist claims).

But, says Strawson, perhaps the personal reference can be eliminated by suggesting that "all my experience" means the same as "all experiences contingently dependent upon body B"; and then replacing the second statement (which is alleged to be synonymous—mean the same)—in the original claim (1). This attempt also fails, however, for such a replacement renders the original statement (claimed by the No-Ownership theorist to be contingently true) analytic. For, taking the original statement ("All my experiences are had \( \&) by body B" and replacing the personal reference "all my experiences" by its alleged meaning—equivalent results in the following clearly analytic statement, "All experiences contingently dependent
on a certain body $B$ are had \(1\) (which "had" means "are causally and contingently dependent upon") by body $B$.

So, once again, the point is that the No-Ownership theorist cannot eliminate the personal reference of the original statement \(1\) which is said to be part of the generation of the illusion concerning the Ego and "had \(2\)".

Given, then, that Strawson has shown that the personal reference (the reference to "my experiences") is a necessary (non-eliminable) part of statement \(1\), which is the first step in the No-Ownership theorists' argument concerning how the illusion of an Ego and "belonging to \(2\)" is generated, Strawson, then, simply points out that the defining characteristic of the class of experiences referred to as mine (my experiences) is that these experiences are my experiences; that is, that they "belong to \(2\)" me. That is, take for example, a pain which I have. It does not seem to make sense to suggest that this very same pain is a pain which you or someone else could have. That is, it seems to be that this pain \((x)\) is uniquely distinguishable from all other pains by being my pain. Or, to put it in terms of belonging to, this pain \((x)\) is uniquely distinguished by "belonging to \(2\)" me. Or, to put it still differently, the defining characteristic of this pain \((x)\) is that it is mine (that is "belongs to \(2\)" me).

Thus, in order for the No-Ownership theorist to show us how the illusion of an Ego and the relationship "belonging to \(2\)" is generated, (by slipping from \(1\) to \(2\)) he must use the very sense
of belonging to (namely "belonging to \( \_2 \)) whose existence he denies. The NT is, thus, inconsistent or incoherent and should be rejected. 6

Critical Comment on 0.1.

It seems to me that Strawson is quite correct here; both with respect to the need for the No-Ownership theorist to use the very sense of possession ("belong to \( \_2 \)) which he denies, and with respect to their being such a sense of possession.

Concerning the latter, it seems clear that if one looks at the conceptual scheme which we do have, one will find a sense of possession ("belonging to \( \_2 \)) which the No-Ownership theorist claims does not exist. My pain cannot be your pain. Put more strongly, it is conceptually impossible that my pain be your pain. Or, to say the same thing, it is a conceptual law or truth that "My pain cannot be your pain" or that "You cannot have my pain". Which is to say that a correct description of our conceptual scheme reveals that "You cannot have my pain." 7 So, there is a sense of "belong to" which is what Strawson calls "logically non-transferable". My pain belongs to me and only me. Logically (or better, conceptually) you cannot have my pain.

As well, it seems to me that a still further objection to the No-Ownership theorist's generation of the illusion can be offered. Namely, the original statement (1) ("All my experiences are had \( \_1 \) by

---

6 Individuals, p. 92.

7 See chapter iii.
body B") is, in fact, not contingently true. It is clearly not offered as a contingently true identity statement. And it is not a contingently true identity statement. Yet, it seems clear that a part of the explication of the meaning of the subject "my experiences" would be that the experiences in question are had by body B. That is, a part of an adequate answer to the question—what do you mean, your experiences (referring to my experiences) would necessarily include reference to their being experiences had by body B. If this is correct, then, a part of the meaning of the subject term "my experiences" is the predicate term, "had by body B", is—under the widely accepted definition of analyticity—analytic. For, the predicate term is merely a part of the full explication of the subject term. Or, to put it somewhat differently, the predicate term is "included in" the subject term. Thus, statement (1) which the No-Ownership theorist claimed to be contingently true is in fact analytic.

0.2. 1. If NT is correct (A), then one never genuinely ascribes states of consciousness to oneself or others (B).

2. Not (B)

Therefore, not (A); NT is not correct. 8

Critical Comment on 0.2.

Strawson is quite correct here. What is it that allegedly makes such ascription improper? Perhaps that it is misleading, that it leads one to suppose that there is a non-material entity that one

8Individuale, pp. 92-93.
is ascribing mental states to? That is not so. There is no reason at all to hold that the mere ascription of states of consciousness to a person leads many to hold that such ascription is ascription to an ego! On the contrary, person-ascription, as such, would lead one more naturally to suppose that ascription is to a person. The mere ascription of states of consciousness to persons does not in any way determine what a person is. At least, there is no evidence that such is the case. The fact (if it is a fact) that persons generally hold that true self-predication is analytic (that a statement such as "All my thoughts are had by me" is logically true) does not in any way suggest the postulation of an Ego which has these experiences. At most, it merely suggests that any person (whatever a person is) has his own experiences. And this is not misleading; indeed, it is true; trivially true! So the prima facie situation stands. Obviously, states of consciousness are properly ascribable to persons.

How well does NT accommodate the relevant phenomena? NT you will recall, denies the existence of any other entity than a body (with respect, that is, to persons). Thus, NT is what might be called a monistic materialistic theory of self. It does not recognize a unique psychophysical entity, a person, (as Strawson does). It does not recognize an ego, a soul, or even a mind (at least so far as Strawson has reconstructed it). A person, then, is and is only, a body. But while a person is a body on NT, a person is not an entity which can—properly speaking—be said to think, to intend, to calculate, to hope, and so on. That is, it is not an entity that can—properly speaking—be said to possess (either possess 1 or possess 2).
states of consciousness. Thus, whatever evidence one has for a
person, that evidence cannot properly involve states of conscious­
ness. With respect to persons doing something, the answer depends
upon whether the "doing" is a "doing" which involves states of con­
sciousness. For, if it does, then a person cannot, properly speak­
ing, ever be said to do anything. For example, in so far as raising
one's arm, or playing chess, or running, and so on, (all "doings")
involve (as they surely do) states of consciousness, then predicating
a "doing" of a person, is also to predicate a state of consciousness.
And, since states of consciousness are never, properly speaking,
predicated of persons, "doing" are never able, properly speaking, to
be predicated of persons NT, then, cannot accommodate PH.1.

With respect to PH.2., 3., and 4. NT would seem to be in a
much better position. For, it does not have the problem of, for
example, CT; namely, the problem of having to re-identify, identify,
and individuate a non-material entity in order to be able to re-
identify, identify and individuate a person. Re-identification,
identification, and individuation can be done by ordinary bodily
criteria (since a person is merely a body). For example, in so far
as a person looks, roughly speaking, the same now as yesterday, we
are, under normal circumstances, justified in believing that he is
the same person. As well, a person is ordinarily successfully iden-
tified by his hair color and style, his skin color, his bodily shape,
his eye color, and so on (his major and minor bodily characteristics).
And, lastly, in normal circumstances, in so far as it "looks like"
(again relying upon bodily (empirical) data) there are three persons over there, then, generally speaking, there are three persons over there.

So, on the whole, NT can accommodate PH.2., 3., and 4. And this fact is surely one of the strongest points in its favor.

However, NT is not completely satisfactory with respect to PH.2., 3., and 4. For, there are circumstances when bodily means of re-identification, identification, and individuation simply are not sufficient for re-identification, identification, and individuation of persons. That is, there are circumstances where mental characteristics play a role in re-identification, identification, and individuation. For example, it has happened, unfortunately, that a person's body was very radically altered as a result of a fire, or car accident, or operation, or whatever. In certain cases of this sort, identification generally shifts from reliance upon bodily characteristics (since this is no longer reliable) to reliance upon non-bodily characteristics, such as what the person says (his memories, thoughts, attitude, desires, goals, ...).

Another example, much more common and plausible, is the meeting of an old school friend whose physical appearance has radically changed (or at any rate, has radically altered in terms of one's memory of that person), and the slow re-identification that takes place ("Yes, he is the same person...") almost totally as a function of non-bodily characteristic recognition. For example, he mentions things you and he did together (memories), goals which he
had (and you remember as being characteristic of him), desires which
he has fulfilled or failed to fulfill (and which were characteristic
of him), and so on.

Now the point of all of this is that while NT can, in most
ordinary circumstances, accommodate PH. 2., 3., and 4., a theory which
could do this, plus accommodate the occasional unusual circumstances
where non-bodily re-identification, identification, and individuation
are relevant, would be preferable even to NT, (in respect, that is,
to PH. 2., 3., and 4.). Strawson's theory can, I will later argue,
accommodate both (and is thus, in respect to these phenomena, prefer-
able).

With respect to PH. 5. (The Phenomenon of Ordinary Language),
NT again has problems. For, of course, many ordinary statements
involving persons involves the predication of states of conscious-
ess to persons. For example, "John is now thinking", "John is now
hoping", "John believes in 'God'", all appear to be instances of
predicating states of consciousness. And quite obviously, statements
like those are often meaningful and true (proper). Yet, according
to NT no such statements are ever proper. Thus, NT cannot accommo-
date PH. 5.

In terms of PROBLEM-SOLVING, NT does at least have a clear and
simple theory of the nature of a person (PR.1.). A person or self
is a body. Indeed, this is one of the major attractions of this
theory. There simply are no mysterious or problematic or "complex"
entities (like an immaterial ego or soul, or a set of perceptions of
various sorts in various relations) involved in the specification of a person. A person is a material body, period.

What, though, about other persons (PR.2.)? Can NT explain how we know that there are persons other than ourself? The answer to this question depends upon the answer to another question; namely, does the existence of a person depend upon the existence of states of consciousness which that person has? Put still differently, if we knew nothing at all about the states of consciousness of an alleged person (not even that they (states of consciousness) existed) would we be justified in saying we knew that that (other) person existed? It seems very obvious that the answer to this latter question is no. If there was no evidence of any sort that P (an alleged person) ever thought, or hoped, or desired, or dreamt, or worried, or...and so on, then obviously one would not be justified in saying he knew that an other person (P) exists. In short, surely one necessary condition of knowing the existence of a person is knowing the existence of at least some states of consciousness (of that person). Yet, according to NT, a person is merely a body. Thus, to know the existence of a person would seem to require knowing only the existence of a body. We have just shown, however, that to know the existence of a person it is not sufficient to know the existence of a body.

NT cannot solve PR.2. for another reason. Any states of consciousness ascription is improper (incorrect). Thus, NT does deny that persons desire, worry, hope, and so on. For, NT denies that such statements as "John desires x" or "John is worried about Jack"
are ever proper (correct, true). But, once again, if to know of the
existence of others, one must know of the truth of some such things
as "Jack worries", "John desires x" and so on (namely, some of his
states of consciousness), then again NT must fail to solve PR.2.

NT also has a unique problem with PR.3. (The Problem of
Solipsism). For, the question here is, how can one (a person) know
that something other than himself exists? But the No-Ownership
Theorist, if he is to be consistent, cannot ever grant that—properly
speaking—we can know that there exist non-person entities. Why?
Because obviously, "knowing" involves, in some way, states of con­
ciousness. Thus, properly speaking, persons are not able to know.
So, once again, the absurdity of denying that persons have (in any
legitimate sense) states of consciousness prevents NT from fully
succeeding as a theory of self.

Lastly, NT not only fails to offer a clarification of the
distinction between the mental and the physical aspects of a person
(PR.4.); it essentially denies that, properly speaking, a person has
any mental aspects. This itself would not necessarily be damaging
to the theory if it could successfully explain why most persons
(wrongly) make a distinction between these two aspects; and why most
persons (wrongly) believe that they genuinely (properly) predicate
mental states of themselves and others. The theory, however, does
not explain this. The theory, of course, also fails (since it does
not even grant that a person has two aspects) to explain the rela­
tionship between these two aspects. In short, NT fails to solve PR.4.
Concerning III (GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES)

and specifically P.1. (The Simplicity Principle) NT meets P.1. perfectly. For NT the self is one thing, a body. Clearly, this is a major attraction of the theory. Unfortunately, it purchases this simplicity at the high price of absurdity.

Again, P.2. remains idle. And, lastly, NT does not violate P.3.

Let us briefly compare the merits and demerits of CT, HT, and NT. All three theories fail with respect to The Problem of Other Minds (PR.2.), The Problem of Solipsism (PR.3.), and The Mind-Body Problem (PR.4.), (though HT and NT, if otherwise acceptable, would not, at any rate, generate the problem as does CT). Thus, it will be especially interesting to see if Strawson's theory can remedy these major failures (Strawson, of course, claims that his theory can). Both CT and HT fail in the very important areas of our successful re-identification (PH.2.), identification (PH.3.), and individuation (PH.4.) of persons. NT (partially) succeeds in these areas because it can rely upon ordinary bodily criteria. As we shall soon see, Strawson will try to retain the success of NT in these areas while remedying the failure of NT in the other areas (of PHENOMENA ACCOMMODATION and PROBLEM SOLVING). Thankfully, all the theories succeed in making clear their view as to the precise nature of a person. Presumably and hopefully, Strawson's theory will also succeed in this area. NT, again, has the edge in relation to Simplicity (P.1.). And, once again, Strawson will try to duplicate the success of NT in this regard. Only CT suffers from ad hoc attempts. But all three have objections over and above the objections
generated by not solving the requisite problems.

So, once again, we see that there has been some historical progress in the Philosophy of Self. For NT is clearly superior to HT and HT is clearly superior to CT (in relation to the structure I have developed). But even NT retains a number of significant failures. Thus, a more successful theory of self is possible and desirable. We shall soon see if Strawson has produced just such a (more) successful theory.
In this chapter, I shall try accurately to describe and clarify Strawson's theory of self. In the next three chapters I shall subject Strawson's theory to the same critical scrutiny as was applied to Strawson's version of Descartes', Hume's, and Wittgenstein's theory. I shall then be in a position rationally to judge the comparative preferability of Strawson's theory.

To begin, let us try to construct an answer to the question, "For Strawson, what is a person?" First, a person is a substance. That is, a person is an entity, a natural entity, "in" space and time. Further, a person is an entity which has certain properties, which has experiences. And lastly, (in relation to the substantial nature of a person) a person is in some sense a permanent, a persistent object. As Strawson puts it (speaking in the context of Kant's theory of self and agreeing with this aspect of Kant's theory),

A man is something perceptibly (if only relatively) permanent, a persistent and identifiable object of intuition, a possible subject of a biography or autobiography.

---

2. *Individuals*, pp. 82-83.
Instead of talking, dubiously, of an experiential route through the world, of one series of experiences constituting such a route, we may talk, confidently, of an undeniably persistent object, a man, who perceptibly traces a physical, spatio-temporal route through the world and to whom a series of experiences may be ascribed with no fear that there is nothing persistent to which they are being ascribed. 3

As well as seeing a person as a substantial entity, Strawson sees a person as the type of entity to which both mental and physical predicates appropriately apply. Put linguistically, the concept of a person is the concept of an entity (a person) such that necessarily both mental and physical predicates are appropriately ascribed and ascribable.

As mentioned in chapter iii, Strawson is here using the term "necessarily" with respect to our conceptual scheme. That is, he can be seen as making the claim that it is a conceptual law or truth that the concept of a person is necessarily the concept of an entity to which both mental and physical predicates are applicable. How did Strawson arrive at this? By observing our language and by observing certain non-linguistic data. He inquired as to the types of predicate which are in fact ascribed to persons. He observed that we do identify, re-identify, and individuate persons and that we use empirical criteria to do this. This suggests that persons must be physical beings of some sort. And the concept of a person must at least be the concept of an entity such that physical predicates are ascribed and ascribable. Yet, clearly, mental predicates are

3 The Bounds of Sense, p. 164.
ascribed to persons. Thus, the concept we have of a person must include the ascription of mental predicates.

In short, a true description of our conceptual scheme reveals the concept of a person to be the concept of an entity to which both mental and physical predicates are ascribed and ascribable. Strawson states this descriptive fact about the conceptual scheme by the phrase "the concept of a person is the concept of an entity such that necessarily both mental and physical predicates are appropriately ascribed and ascribable." Similarly, the following is the way a descriptive fact about the physical world is often put, "If an unsupported body is dropped, it must (it necessarily will) fall to the ground."

Strawson distinguishes more specifically between the two types of predicates which he designates M-predicates and P-predicates. M-predicates are such predicates as "weighs 10 stone", "is in the drawing room" and so on. These are predicates which are properly applied to persons; yet are also properly applied to non-persons (other material bodies), "to which we would not dream of applying predicates ascribing states of consciousness." On the other hand, P-predicates are such predicates as "is smiling", "is thinking hard", "believes in God", and so on. This, says Strawson, is the class of all predicates which apply to persons, other than M-predicates.

^Individus, pp. 100; 132.

^Individus, p. 100.
Strawson also describes P-predicates as having in common the implication of "the possession of consciousness on the part of that to which they are ascribed."\(^6\)

It is, thus, not the case that P-predicates are predicates properly ascribable to persons only. They are properly ascribable to any entity which possesses "consciousness." Thus, all P-predicates are not necessarily "predicates ascribing states of consciousness", but their proper ascription does imply that that to which they are ascribed does possess consciousness.\(^7\) As an example of a P-predicate which is not a state of consciousness predicate, Strawson offers the predicate "going for a walk". The concept of a person, then, (in terms of P-predicates and M-predicates) is the concept of an entity to which both M-predicates and P-predicates are properly ascribed and ascribable.

A still further characterization of the concept of a person is offered by Strawson. Strawson claims that the concept of a person is "logically primitive." What does this mean? Strawson first tries to tell us what it excludes with respect to the person-concept. It means first that the concept is not to be thought of as a secondary concept, to be understood in terms of two primary concepts, one of a particular conscious, and two, of a particular human body. That is, the concept of a person is not to be thought of as the concept of a body and a mind (as with CT). Put still differently, the concept of

\(^6\) Individuals, pp. 100, 101.

\(^7\) Ibid.
a person (John) is not to be thought of as type-ambiguous. That is, it is not to be thought that "John is heavy" and "John thinks" are two sentences such that John refers to (denotes) one thing in the first sentence, and John refers to (denotes) a different thing in the second sentence. Thus, to say the concept of a person (John) is logically primitive is at least to affirm the above negations.²

Another explanation of the logical primitiveness of the concept of a person is that the concept of a person is logically prior to that of, for example, an individual consciousness. Apparently what Strawson means here is that if one had no concept of a person, one could have no concept of an individual consciousness. That is, it is not the case that one has a concept of an individual consciousness, and a concept of an individual body, then from these one generates the derivative (secondary) concept of a person; rather, one has a concept of a person (an entity to which both M-predicates and P-predicates properly apply) and from this concept is generated the derivative (secondary) concept of an individual consciousness (of that person) and of an individual body (of that person). But where does the logical primitiveness come in?

I don't believe that Strawson is asserting that it is, in the usual sense (i.e., such as to generate a contradiction), logically impossible that the concept of a person arise from the concept of an individual consciousness and the concept of an individual human body.

²Ibid.
But he is asserting that, given our present conceptual structure (the way we refer to persons, our successfully identifying, re-identifying, and individuating persons, and so on), which conceptual structure Strawson is intent upon describing (he is, one must remember, concerned with the "logic" (logical geography) of our concepts), the concept of a person is "logically" primitive. That is, the "logic" of our conceptual scheme is such that the concept of a person is primitive. Translated, that should mean, given the way we refer to persons, the way we identify, re-identify, individuate, and so on, persons (in short, the way we wield the concept of a person) the concept of an individual consciousness is generated from the concept of a person, and not vice versa. In that sense, the concept of a person is primary with respect to the concept of an individual consciousness; and the concept of an individual consciousness is derivative (secondary). Says Strawson, "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." Now in this quotation, I take Strawson as meaning, when he says "cannot exist", as meaning logically cannot, but logically cannot in the

9 See chapter ii and iii.
10 Individuals, p. 99.
sense that—given our conceptual scheme, cannot—rather than logically
(in the usual sense, "such as to generate a logical contradiction")
cannot. I so interpret Strawson because (1) there is no reason to
believe that in the usual sense of logical, such a concept (of an
individual consciousness) could not be logically primitive (for
example, with respect to some different conceptual scheme); and (2)
Strawson does claim to be describing our conceptual scheme; not any
conceptual scheme. One must, then, be very careful to distinguish
Strawson's uses of the term "logical". In short, that the concept of
a person is "logically" primitive is a conceptual law or truth. 11

In sum then, to say that the concept of a person is logically
primitive while the concept of an individual consciousness is logi-
cally derivative or secondary, is to say that—given our conceptual
scheme (given the way persons wield the concepts making up our con-
ceptual scheme) the concept of an individual consciousness would not
be meaningful or useful, would not exist, if there were no concept of
a person (who has the consciousness). Put somewhat differently, the
"logic" of our conceptual scheme (which Strawson purports to be
describing) is such that the concept of an individual consciousness
does derive from our concept of a person.

Very closely related to Strawson's viewing the concept of a
person as logically primitive is his view of a person as a unique

11 See chapters ii and iii.
ontological entity. That is, it is clear that Strawson does not view a person as solely a mental entity (for bodily predicates are properly applicable to persons). Nor does Strawson view a person as solely a bodily entity (for mental-predicates are properly applicable to persons). And yet neither is Strawson willing to view a person as a two-thing entity; that is, as a mind and a body. For, this is the Cartesian view, which Strawson clearly rejects. As well, Strawson says clearly that "The concept of a person is not to be analysed as that of an animated body or of an embodied anima."¹² That would seem to be a linguistic way of saying that a person is not to be viewed as a body with a soul (or mind); or as a soul which has a body. Thus, put non-linguistically, a person is not an ontological conjunction. A person is a unique ontological entity. Put linguistically, for Strawson, the category of persons is one of the irreducible categories of the world.¹³ A correct categorical description of this world would necessarily include the category (concept) of a person.

The last important aspect of Strawson's theory is its stress on the agency character of personhood. Says Strawson, "... the idea of oneself as an agent forms a great part of the idea of oneself." Indeed, Strawson even suggests that this agency aspect is "perhaps a

¹² Individuals, p. 99.

¹³ Some examples of reducible categories would be: Hume's concept of a person as reducible to perceptions and relations; Descartes' concept of a person as reducible to a mind and a body; Hume's concept of causality as reducible to perceptions and constant conjunction.
necessary part" of the idea of oneself.\footnote{x} And lastly, Strawson observes 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."\footnote{y} Thus, clearly another way of describing Strawson's theory is that for Strawson a person is a natural object, with a certain nature, which \underline{does} things and has things done to it. Just what the various characteristics are which make up this "common human nature" Strawson does not say.

The above, then, is a full sketch of Strawson's person-theory. Before moving on to the next chapter, it might be interesting to see Strawson's theory in historical perspective. How \underline{different}, how \underline{original} is Strawson's theory? Is it merely a resurrection of an earlier theory? The two theories which may be similar to Strawson's are the theories of Aristotle and of Spinoza. Is Strawson's theory but a resurrection of either or both of these theories? I shall examine each in turn.

For Aristotle, a person is a substance with a nature. But Aristotle's notion of (Primary) Substance includes the following:

\( x \) is a (Primary) Substance if

1. \( x \) is capable of admitting contrary qualities.
2. \( x \) has no contrary.
3. \( x \) does not admit of variation in degree.
4. \( x \) "underlies" other things.

\footnote{x}Individuals, p. 77.
\footnote{y}Individuals, p. 109.
5. x is not predicatable of a primary substance (a subject) and is capable of existence apart from a primary substance. 
6. x is an individual.  

As well, for Aristotle a person is a body and a soul; but the soul is simply seen as a set of dispositions and capacities and actions (thinking, wishing,...) of a body. So it is not the case that a person is two things. A person is a body with certain capacities and which acts in certain ways. On the whole, then, the soul (mind) is not separable from the body. It must be mentioned, however, that Aristotle did recognize one "part" of the soul which is separable from the body and which survives the death of the body; this is the Active Intellect.

Comparing Aristotle's theory to Strawson's, one can see that there is some similarity: For example, both hold a substantival view of the nature of a person (as opposed, for example, to a process view, such as James' or an event view, such as Hume's). Both to some extent deny that a person is made up of two ontologically distinct, separable entities. Strawson holds this unequivocally; Aristotle holds it except with respect to the Active Intellect. Both stress the bodily nature of a person; man is very much seen as a natural object among natural objects.

There are, however, many major differences in their theories. First, Aristotle has a considerably richer conception of the substantival nature of a person than does Strawson. That is, Aristotle

16 See Metaphysics, pp. 805, 785, 761; Categories, 9, 10, 12, 13, 14; in The Basic Works of Aristotle, edited by Richard McKeon (Random House, New York, 1941).
sets out an extended set of substance-making characteristics. Strawson does not use Aristotle's conception of substance; indeed, he does not use the term at all. But he clearly adopts a kind of substance view in one ordinary use of the term "substance". He views persons as more or less (persons do perish eventually) permanent individual entities "in" space and time, which have certain properties. This, however, is far from (though not necessarily incompatible with) the use of Substance adopted by Aristotle.

Next, Strawson stresses much more than Aristotle certain linguistic features with respect to persons; for example, the primitive character of the concept of a person; and, the distinction between M-predicates and P-predicates. Further, Strawson distinguishes a special and separate person-category. Aristotle does not. For Aristotle, persons fall into the broad category of Substance. Lastly, Strawson's theory lays much stress on the agency character of a person. Aristotle's theory does not (though Aristotle's theory does affirm the agency character of a person). Thus, Strawson's theory is not simply a resurrection of Aristotle's theory; though there are some similarities.

Spinoza also developed a monistic theory of persons. And once again a concept of Substance plays a crucial role in the theory. For Spinoza, a person is a part of Substance. And Substance is conceived as "that which is in itself and conceived through itself."

---

There is one and only one Substance; that is the universe, also called Nature, also called God. However, this one Substance manifests itself to humans in at least two ways; either as Thought, or as Matter (Extension). The One Substance manifests itself via two major attributes (essential properties of the Substance); Thought and Extension. Thus, for Spinoza as well as for Strawson, a person is one thing; not two different things. Says Spinoza, "the idea of body and body, that is, mind and body . . . are one and the same individual conceived now under the attribute of thought, now under the attribute of extension . . ."\(^{18}\)

Once again there is some similarity to Strawson's theory. Both reject the Cartesian view of a person. That is, both are monistic theories. Both view a person as one thing; though both recognize in their respective ways the mental and physical aspects of a person. But while there is this basic similarity, Strawson's theory goes far beyond and is thus far different from the theory of Spinoza. For example, Strawson's theory stresses certain linguistic aspects of the concept of a person; such as the defining function of the predication of M-predicates and P-predicates; also, the primitiveness of the concept of a person. Furthermore, Spinoza's theory is closely bound up with a notion of Substance which Strawson does not make use of. Still further, Strawson recognizes the category of person as a basic ontological category. That is, he does not subsume the category of Person under the very broad basic category of

\(^{18}\text{Ibid.}, \text{p. 102.}\)
Substance (God). And lastly, Strawson's theory stresses the agency character of personhood. Spinoza's theory does not.

Strawson has, then, produced a relatively original and surely interesting theory of persons. In the next few chapters I shall attempt to discover if Strawson's theory is not only interesting but also correct (that is, preferable to the other major theories which I have critically examined).
CHAPTER VIII

STRAWSON'S THEORY AND PHENOMENA ACCOMMODATION

This chapter will initiate the critical examination of Strawson's Theory (hereinafter referred to as ST). How well does ST accommodate the phenomena (PH.1. - PH.5.)?

PH.1. First, Strawson makes it very clear that he does recognize this phenomenon. He, of course, admits that persons do things and have things done to them. He further suggests that our concept of a person is largely a concept of an agent. He observes that in terms of our actual conceptual scheme, we do indeed appropriately attribute to ourselves and others, doings. He even suggests that we necessarily do so. ¹ But a concept of oneself does not logically (in the usual sense) necessitate a concept of oneself as agent. One presumably could develop a conceptual scheme where one had a concept of oneself (even as distinguishable from other non-self things) which did not involve oneself as agent (and do so without contradiction). But once again Strawson is using the term "necessary" in relation to a broader notion of logical; namely, in relation to the "logic" of our conceptual scheme. And he means that --given our conceptual scheme as it is, the concept of oneself does

¹*Individuals*, pp. 77, 78.
involve the concept of agency. And to say this latter is to say that an accurate description of the concept of oneself as it occurs in our conceptual scheme will include its agency aspect. We, therefore, have another conceptual law or truth.

So, ST does recognize PH.1. And, from a linguistic point of view, ST has no difficulty with the agency aspect of personhood. For, the concept of a person is specified by Strawson as the concept of an entity to which both M-predicates and P-predicates appropriately apply. Now if agency predicates are either M-predicates or P-predicates, then, of course, agency predicates are appropriately applied to persons. But in order to decide if they are either M-predicates or P-predicates we must ask what exactly Strawson means by agency? With respect to the possibility of securing an agency aspect to a self as conceived of in a No-Space (purely auditory) world, Strawson writes,

We have introduced no distinction between moving and being moved. Suppose we introduce such a distinction. Suppose, that is to say, that the being whose experience is purely auditory sometimes just suffers change of position--change just occurs--and sometimes initiates it. (If anyone asks how this is to be understood in terms of movement along an auditory scale, I refer him to differences in the way he anticipates what he is going to do and what is going to happen to him--differences in the kinds of knowledge he has of these two things.) It might seem that the introduction into our universe of this distinction--the distinction, roughly speaking, between changes that are brought about, and changes that merely occur--could necessitate the introduction of the idea of that which brings about the deliberate changes,....

[2] Individuals, p. 76.
It would seem, then, that an agent is an entity which can be truly said to "bring about change". But Strawson also seems to suggest that intention is involved in agency. Let us, then, interpret Strawson's concept of agency as involving: an entity which brings about change by way of intending. (On this interpretation, since intending implies consciousness on the part of the intender, agency predicates are P-predicates.)

Now the question arises, can ST accommodate PH.1.? If agency means "to bring about change by way of having an intention" that would seem to imply that some sort of a mental component is clearly involved. Does this produce difficulties for Strawson? I believe so. For, Strawson must satisfactorily answer the question (1) How can a mental occurrence play a role in producing (causing?) a physical occurrence (a change)? And indeed, this question has a preliminary: namely, (2) What is it that distinguishes mental characteristics from physical characteristics? That is, the answer to question (1) may depend upon the answer to question (2). Does Strawson answer or even try to answer either of these very obvious and crucial questions? He does not. And, this is one of the weakest aspects of Strawson's philosophy of mind. What he does is merely distinguish two types of predicates: M-predicates and P-predicates. M-predicates are what we have called physical predicates. P-predicates are what have been called mental predicates.

---

3Ibid., pp. 77, 108.
Strawson does, thus, recognize and grant a distinction between mental and physical, but merely on a linguistic level. What is the nature of the reality behind the differing predicates? The most we get is that the reality behind the physical predicates is physical. And the reality behind the mental predicates "imply the possession of consciousness on the part of that to which they are ascribed."\(^4\)

But this does not help at all. For, first, we are still not told what mental characteristics are. We are merely told that those things which possess mental characteristics also possess consciousness (this would seem to be the force of the implication above). Thus, mental characteristics are not said to be identical with consciousness.

But even if it were said that mental characteristics are conscious characteristics, that would still not be unproblematic. For, it is not at all clear what consciousness is! Incredible as it may seem, no attempt at all is made by Strawson to deal with this most crucial question: what distinguishes mental characteristics from physical characteristics? To say that mental characteristics are conscious and physical are not, is simply to generate a rephrase of the original question: "What distinguishes conscious characteristics from non-conscious (physical) characteristics?" One has gained little (if anything) from such a rephrase.

The point of the above is that Strawson simply cannot plausibly explain how a person can be an agent if being an agent involves mental characteristics.\(^4\) *Individuals*, p. 101.
and physical characteristics and a relation between them and yet no satisfactory definition or understanding of mental and physical characteristics is offered.

Strawson's problem is thus one of what might be called ontological explanation. He is here apparently unwilling to move from the linguistic or conceptual level (talk of M-predicates and P-predicates) to an ontological level (talk of the referents of these predicates). It is one thing to say that a Strawsonian concept of person (subject) can be appropriately, sometimes truly coupled with agency predicates. It is another to say that one can explain how a Strawsonian person can do this or that. With respect to the former, Strawson's theory is successful. With respect to the latter, it is not.

PH.2. Clearly if one views a person as a non-natural, non-bodily object, one is going to have difficulties explaining how persons are, as they are, re-identified. But if one views a person as a natural object, a necessarily bodily object, one can then simply appeal to ordinary empirical criteria to explain re-identification and the unity or identity of a person through time. This is the strength of Strawson's view. Says Strawson (speaking of Kant but clearly in agreement with him), "If we are to make any legitimate employment of the crucial concepts of unity or numerical identity through time, we must apply them, in the light of empirical criteria, to objects encountered in experience." A Strawsonian person is an

5The Bounds of Sense, p. 37.
object encountered in experience.

Strawson does not, however, offer any examples of some of these empirical criteria which we presumably use. But, I shall try to give some examples of what Strawson might have in mind. John and Jack are friends. They are together. They part. Next day, John says, "Hello Jack." A re-identification has taken place. How? How did John re-identify Jack?

A whole set of empirical data plays a role. For example, the following set would clearly justify (make highly probable, though not certain) calling person P (the person John says "Hello" to) "Jack":

1. P looks like Jack. (Same eyes; nose; facial features, color hair, height, skin color, and so on.)
2. P talks like Jack (same voice, tone, inflections, and so on.)
3. P acts like Jack (empirical data).
4. P thinks like Jack (as evidenced by the things P says (empirical data)).

The major point is that a person Jack can be re-identified by one’s knowing whether or not certain sets of M-predicates (such as, is tall, has red hair, has a scar on his left hand, and so on) and P-predicates (such as, believes in God, remembers the day you and he caught a three-foot catfish, and so on) which distinguish that particular person Jack, are true of the person you now meet. If they are and if there is a general continuity with respect to new predications (new thoughts, actions, and so on) then that person you now meet is Jack.  

I am here (with respect to new predications) going beyond anything Strawson explicitly says. I believe, however, it is necessary to do so, to better accommodate our actual conceptual scheme.
In short, then, what is it that constitutes the sameness (the unity) of Jack? A stable set of M-predicates and P-predicates; and, a general continuity with respect to most new predications. By the latter I mean merely that new predications (such as those having to do with Jack's actions, and thoughts) should have a number of clear relationships with past predications. This latter is important; for one is often tempted to say of a person (when this general continuity seems to break down), "You know, Jack really is a 'different person' now." And, of course, there is a sense of different person where to say this is meaningful and true. But when pressed as to whether he is "really" a different person (now in a different and more fundamental sense), the speaker would probably admit that he is still the same person. Here the criterion shifts from the new predications back to the past predications. Or the criteria shift might be from mental predications (emphasis upon his change in ideas) to physical predication (emphasis upon how he looks, sounds, and so on). Thus, interpreted and expanded (as I have expanded it), ST can accommodate PH.2.

PH.3. The same general sorts of considerations apply to PH.3. We identify Jack as Jack by knowing that a particular set of M-predicates and P-predicates truly apply to Jack. And Strawson can make direct appeal to empirical criteria for identification. What might be an example of identification? You are in a three-man conversation. You, Jack, and John are involved. You identify Jack as Jack, and John as John, on the basis mainly of two sets of empirical clues available to you: 1. the bodily features and 2. the sounds made
(the tone of the voices, the content of the conversation--memories, thoughts revealed, and so on). You do not have to be seen as identifying a mind (Jack's mind) and independently identifying a body (Jack's body). Rather, you are identifying a person; an entity which has both bodily aspects and mental aspects. There is, for Strawson, no wedge between these two aspects of a person such as to make a person a two-entity thing. And thus Strawson avoids the problem arising from having independently to identify two entities before identifying a person. In sum, identification is basically of persons, not of bodies and of minds. And ST escapes the problems which we examined with respect to independent mind-identification. ST can, thus, quite readily accommodate our having "clear, ordinary, and generally successful (though not infallible) ways of identifying persons."

PH.I. ST likewise escapes the problem of individuation. For, again it does not have to explain how individuation of minds can and does occur independently of individuation of bodies. This, you will remember, is what the Cartesian theorist (as interpreted by Strawson) had to do. Rather, Strawson can simply point to how persons do in fact individuate persons. They use clear, ordinary, empirical criteria. For example, (once again the examples are mine--for, strangely, Strawson never offers even one example of these empirical criteria which he relies so heavily upon), take a group of four persons which

---

7I use the term "bodily aspects" as a summary term to refer to the various referents of M-predicates. Likewise, the term "mental aspects" refers to the various referents of P-predicates.
are observed. How does one know that there are four persons in a room? Clearly there are cases where one does know that there are four persons in a room. Thus, there must be some criteria (whether implicit or explicit) which we use to correctly judge that there are four persons in a room.

I wish to suggest that, not surprisingly, a large part of our criteria is visual. That is, a large part of our justification for asserting that there are four persons in a room is that we see four and only four persons in a room. This criterion, namely, a visual criterion, does make sense with respect to Strawson's person, for one can see a person (an entity to which both mental and physical predicates appropriately apply). It does not make sense with respect to Descartes' person (as interpreted by Strawson); for, one sees a body (using "see" in its ordinary physical sense), but one does not see a mind. But a person is not just a body; for Descartes he is also a mind. Thus one cannot correctly be said to see a person. One sees a human body; one infers that there is also a person there.

There are other empirical criteria of which we make use in conjunction with the visual criterion. For example, an auditory criterion is often used as well. The fact that four and only four different voices come from the room would be a relevant consideration in deciding how many persons were in the room.

These two criteria are surely in most cases the major empirical criteria used for individuation. However, other empirical criteria can be and sometimes are used. For example, an especially sensitive
person (for example, a blind person) could use an olfactory criterion to individuate the four persons. Likewise a person could use a kinesthetic criterion to individuate the persons. And last, and least likely but still quite possibly a person could use a gustatory criterion to individuate the persons.

The point of all this is first, that persons do in fact use these empirical criteria (especially the visual and less often, the auditory) to individuate persons; and secondly, Strawson's view of what a person is enables him to grant that individuation of persons can and does occur and does so via such empirical criteria.

Before examining ST and PH.5, it might be helpful to summarize and further clarify why ST can accommodate PH.2 thru PH.6, while, for example CT cannot. For CT, to re-identify, identify, and individuate a person, one must re-identify, identify, and individuate two entities: one, a mind; and two, a body. For, a person is not just a body, nor just a mind. But we have argued at length that we have no way to re-identify, identify, or individuate minds. Therefore, CT cannot re-identify, identify, or individuate persons.

But on ST, re-identification, identification, and individuation can take place without difficulty. For, a Strawsonian person is not two entities but one. And, one can use ordinary empirical criteria (methods or ways of telling) to re-identify, identify, and individuate. That is, for example, in so far as one is re-identifying an entity such that it has both corporeal and mental attributes, one can use empirical information to determine which corporeal and which mental
attributes any given person has. How would you know that an entity is truly said to have red hair (an M-predicate)? You look and see if the entity being referred to has red hair. How would you know that this same entity is truly said to believe in God (a P-predicate)? You do the things necessary to provide good evidence that he does. You ask him if he does. You check out his reliability (does he tell the truth about these matters). You observe his behavior. You ask his friends. And so on. (All empirically accessible information).

What the above reveals (and we shall examine this in even more detail, later) is that the logic of P-predicates is such that they are ascribable to other persons on a basis different from their ascription to oneself. That is, one can know directly (that is, without asking other persons, without observing one's own behavior, without checking out one's own reliability, and so on) that one, for example, believes in God. So, true self-ascription does not rely (as does true other-ascription) on observations of the above sort (of one's behavior, of one's reliability, and so on). In sum, then, there is no problem about using empirical information to truly predicate M-predicates (since physical characteristics are obviously empirically accessible). Secondly, self-ascription of P-predicates is not mysterious and is not questioned (even by the sceptic) and is clearly a phenomenon. And thirdly, the logic of P-predicates is such that they are correctly other-ascrivable on the basis of observation. That is, it is a conceptual truth that P-predicates are other-ascrivable on the basis of observation. Thus, P-predication also relies upon empirical information, empirical criteria.
Since a person is an entity such that necessarily both P-predicates and M-predicates apply, and since both such types of predicates are, in fact, truly ascribable on the basis of empirical information and criteria, a Strawsonian person can be re-identified, identified, and individuated.

PH.5. If the concept of a person is interpreted as the concept of an entity such that both mental and physical predicates appropriately apply, then of course predictions such as "John is now thinking", "John is now walking", and so on, can be granted to be appropriate and sometimes true. And such a predication as "John is a person; not just a thing." is quite appropriate and meaningful in terms of ST. For, a person is considered a unique ontological type for Strawson. So once again Strawson's theory enables one to grant unproblematically that all sorts of statements which persons make about persons and which persons believe to be meaningful (not just tautologies) and true, are, in fact, meaningful and sometimes true. And clearly, in so far as other theories (such as CT, HT, and NT) fail to be able to grant this they are less preferable as theories. (Unless, of course, they can show these phenomena to be illusory.)
CHAPTER IX

STRAWSON'S THEORY AND PROBLEMS TO BE SOLVED

In this chapter I shall examine ST in relation to II. (PROBLEMS TO BE SOLVED).

PR.1. In chapter vii, a very thorough presentation of Strawson's conception of the nature of a person was offered. I shall here summarize that part of the presentation which referred to persons as opposed to the concept of a person (and also make one addition to it). For Strawson, a person is an (one) entity (a natural object) "in" space and time. It is an entity which has certain properties. It is a relatively permanent type of entity. Furthermore, it does things and things are done to it. Likewise, a person is an ontologically basic type of thing.¹ This does not mean either that persons are the only things which exist or are real, nor does it mean that everything else can be somehow reduced to persons. Rather, to say "x is basic" means that in order to identify certain types of particulars (y) one must do so by referring to some other type of particular (x), the latter then being referred to as more "basic".² For example, private experiences, private particulars (y) (such as 'my thought now' or

¹ Individuals, p. 256.
² Ibid., p. 4.
'your feeling of love, yesterday') are, says Strawson, an example of a type of particular which is identification-dependent. In order to identify a private experience one must refer to the person whose experience it is.\(^3\) Says Strawson, "A twinge of toothache or a private impression of red cannot in general be identified in our common language except as the twinge which such-and-such an identified person suffered or is suffering, the impression which such-and-such an identified person had or is having."\(^4\) In the above sense, then, persons are an ontologically basic type of entity.

What sorts of answers to the question, "What is a person?" does Strawson explicitly reject? He rejects the view that a person is a two-substance entity. Specifically, he rejects the view that a person is a body and a mind (CT). He rejects the view that a person is a set of private particulars in relation (HT). He rejects the view that a person is merely a body about which it is never proper to say it (the person; the body) is "now thinking" or "has certain feelings" and so on (NT). He rejects all of these views on the basis of arguments and claims which I have, in former chapters, critically examined. I have also suggested that Strawson's theory is not identical to the theory of self of Aristotle, nor to the theory of self of Spinoza; though I did admit that there were some similarities.

With respect to The Problem of Other Persons (PR.2.) it must be observed that Strawson's approach is to argue that the problem

\(^3\)Ibid., pp. 31, 32.

\(^4\)Ibid., p. 31.
cannot plausibly, coherently, be generated. Says Strawson (discussing this problem), "These remarks are not intended to suggest how the 'problem of other minds' could be solved, or our beliefs about others given a general philosophical 'justification'. I have already argued that such a 'solution' or 'justification' is impossible, that the demand for it cannot be coherently stated."\(^5\) So, Strawson will try to dissolve the problem by showing that with a proper understanding (description) of the logic of our language, one can see that the demand for a 'solution' is ridiculous, for the problem itself cannot coherently arise.

Strawson has a major argument with respect to this problem. But before presenting it, we must get clear on a number of points. First, we must clarify a phrase used by Strawson which is essential to the argument, and which has been the subject of much controversy and negative response to Strawson's argument.\(^6\) The phrase is "logically adequate criteria". Unfortunately, the concept of a criterion or of criteria has been the subject of much discussion recently, but there is no general agreement on the subject. However, Anthony Kenny in an article called "Criterion" in the Encyclopedia of Philosophy distinguishes at least three interpretations of the term. One view is that a "criterion for x" is "something which provides decisive

\(^5\) Individuals, p. 109.

evidence or necessarily conclusive evidence for x." A second view is that a "criterion for x" is "something which provides the best possible evidence for x." A third view is the following: "We may characterize the criteria for the truth of a judgment as those states of affairs that are (whose existence would be) direct and noninductive evidence in favor of the truth of the judgment." I shall use these interpretations merely as a guide to Strawson's meaning.

But first, how does a criterion for x get established? That is, let us say the presence of Y is the criterion (using the first of the above interpretations) of the presence of x. How did Y get to be the criterion for x? Wittgenstein suggests that we fix the criterion, perhaps by 1) definition, or 2) by the process of teaching and learning the uses of the terms involved.

Now, since Strawson uses the phrase "logically adequate criteria" it seems plausible to suggest that his use might be similar to or identical with the first view of criteria; for the first does involve necessarily conclusive evidence (suggesting a logical relationship, perhaps a deductive one). However, Strawson does admit that a person could, for example, disguise or fake certain behavior (for example, pain behavior). So, pain behavior cannot itself provide necessarily conclusive evidence for the existence of pain in a person. Perhaps it is something like views two or three that Strawson has in mind?

---


I believe what he means is this. In our conceptual scheme (this is part of the force of "logically", referring to the logic of our concepts) pain behavior (for example) is, in fact, treated as good evidence for the existence of pain. That is, there is "logically" (with respect to the logic of our language) an evidential relationship between behavior and, for example, pain or depression (or whatever). This relationship is established by convention; that is, we learn it along with learning the use of terms like "pain", "depression", and so on. It is established (as Wittgenstein suggests) by learning the language.

It can be added that, ultimately, the strength of the evidential relationship is, of course, dependent upon the circumstances. For example, if a person is screaming he is in pain, and is sweating, bleeding, and in convulsions; and there is no evidence a play is going on, nor a trick is being pulled, and so on, then, the evidence that this person's behavior provides is very strong indeed. Indeed, in ordinary language I believe it could meaningfully be called, conclusive. If, on the other hand, a person's behavior indicates he is in pain, but nothing else does, and you have reason to believe that he is trying to trick or deceive you, then the evidential relation is and should be less compelling. But the important thing is that pain behavior does in our conceptual scheme generally provide good evidence that a person is in pain.

I shall, then, interpret Strawson's use of the phrase "logically adequate criteria" as follows: to say, for example, that the existence of Y is a "logically adequate criterion" for the existence
of x is to say that "the existence of Y is good evidence for the existence of x."

Second point, it should be observed that the concept of criterion (as just explained) and the concept of meaning are not to be considered identical. That is, Strawson wishes to hold that from the fact that, for example, Y is the criterion for x it does not follow that Y is the meaning of x. Says Strawson,

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. So with the language in which we ascribe P-predicates. To say that the criteria on the strength of which we ascribe P-predicates to others are of a logically adequate kind for this ascription, is not to say that all there is to the ascriptive meaning of these predicates is these criteria. To say this is to forget that they are P-predicates, to forget the rest of the language-structure to which they belong.9

Third point, Strawson distinguishes between "Y's being a sign of x" and "Y's being a logically adequate criterion for x." The latter has been defined. By the former, Strawson means "Y provides evidence for x but evidence which is not established by definition or learning and teaching the uses of the terms involved in Y and x, but rather is established by observed correlations between Y and x."10

Fourth point, Strawson believes that the nature of P-predicates

9 Individuals, p. 107.
10 Ibid., p. 102.
(the logic of mental concepts) is such that sets of behavioral criteria (what John says and does) provide "logically adequate criteria" for the appropriate and correct predication of P-predicates. That is, Strawson does not believe that behavior-criteria are merely signs of the presence of states of consciousness.\[11\]

With the foregoing points in mind, I shall now attempt to construct and make plausible Strawson's major argument. Once again, the purport of the argument is to justify our belief that there are other persons by showing that the alleged problem concerning our knowledge of other persons cannot coherently arise.

Or, to put the strategy in my terms: first, it is a phenomenon that we know that there are other persons. The sceptic surely grants that it seems obviously true that we know that there exist other persons. But what the sceptic does is try to show that it is an illusory phenomenon. He tries this by trying to produce a good argument (a good reason) for holding that our alleged knowledge is illusory. What Strawson will try to do is, show that (argue that) the sceptic cannot generate his argument and, thus, has not shown that our alleged knowledge is illusory (and thus we are justified in treating this phenomenon as a fact--namely, that we do know that there are other persons). Once having shown the above, Strawson will then explain how a person can know of the existence of other persons.

\[11\]Ibid., pp. 102-103.
Argument: Part I. The Generation of the Sceptical Problem

1. The relation between the ways of telling if a person has a P-predicate and his having a P-predicate is a sign relation (the ways of telling are signs of the presence of P-predicates.)

2. But, one could only know that A (a set of behavioral ways of telling) was a sign of the presence of B (a mental state; a P-predicate) if one observed correlations between the presence of A and the presence of B.

3. But one can observe this correlation only with respect to oneself.

4. Thus, with respect to others, one has no logically adequate criteria for knowing of the existence of the mental states of others.

Therefore, one cannot know that there are other minds (and thus other persons).\(^{12}\)

Argument: Part II. The Sceptical Argument (Part I.) Cannot Be Generated

1. Self-Ascription (of P-predicates) does sometimes meaningfully take place. \((A)\)

2. Self-Ascription (of P-predicates) meaningfully (usefully) takes place \((A)\), only if other-ascription meaningfully (usefully) takes place \((B)\).

3. If other-ascriptions meaningfully (usefully) take place \((B)\), some ways of telling must constitute logically adequate criteria for the true or correct ascription of P-predicates to others \((C)\).

4. If some ways of telling constitute logically adequate criteria for the true or correct ascription of P-predicates to others \((C)\), then the sceptical problem cannot be generated \((D)\).

Therefore, \((D)\) the sceptical problem cannot be generated.\(^{13}\)

\(^{12}\)This sceptical conclusion also requires the rejection of the ability of an analogical argument to solve the problem. But we have already admitted the failure of the analogical argument.

\(^{13}\)I have reconstructed this argument from passages on pages 102, 103, 106, and 107 of Individuals.
General Defense of Strawson's Argument:

First, I wish to make some observations.

OB.1. Part I. of Strawson's argument merely states (in Strawson's terms) how scepticism concerning the existence of other persons has, in fact, been generated. It is a fair statement of that generation.\(^{14}\) OB.2. It should be mentioned that the usual way of trying to solve this problem is by some form of analogical argument. But, there has been and is much controversy as to whether any form of analogical argument is effective at solving the sceptical problem. A most recent essay suggests that the majority of philosophers find the analogical argument inadequate as a solution to the problem.\(^{15}\)

Now, concerning Part II. of Strawson's argument, I shall first try to suggest what defense, if any, is given for each premise.

Premise 1: This premise is, I believe, obviously true. Moreover, the sceptic who generates the problem of other minds agrees


\(^{16}\)J. M. Shorter, "Other Minds", The Encyclopedia of Philosophy, Vol. 6, pp. 7 ff.
that it is true. He does not find self-ascription of mental predicates to be problematic. For, we have direct access to ourselves and our mental states. It is only ascription of mental states to others which he finds problematic. Thus, Premise 1. is true and is agreed to be true by the sceptic.

Premise 2: Strawson affirms that this premise is true because of the meaning of the terms "you" and "others". He says at various places throughout Individuals:

1. There is no sense in the idea of ascribing states of consciousness to oneself, or at all, unless the ascriber already knows how to ascribe at least some states of consciousness to others. So he cannot argue in general 'from his own case' to conclusions about how to do this, for unless he already knows how to do this, he has no conception of his own case, or any case, i.e. any subject of experiences.

2. ...if only mine, then not mine at all.

3. One can ascribe states of consciousness to oneself only if one can ascribe them to others.17

Now I take it he means by all of this that self-referential terms simply are meaningful (have a use) because and only because other-referential terms are meaningful (have a use). Or, to put it somewhat differently, the concept of oneself (I, Me, ...) is meaningful only because the concept of others (you, him, them,...) is meaningful. Whether this premise is defensible, so interpreted, I shall discuss in the next section (Critical Examination of Strawson's Argument).

17Individuals, pp. 102, 106, and 96.
Premise 3: This premise follows from Part I. and the failure of the analogical attempt to solve the problem. That is, if no ways of telling if other persons have pain, or are depressed, or (and so on, for all other-ascriptions) constitute logically adequate criteria, then there is not good evidence that other persons even exist. But if there is no good evidence that other persons exist then it surely cannot be meaningful to ascribe states of consciousness to others.

Premise 4: The sceptic would accept this premise as true, also. For, of course, if some ways of telling that persons other than oneself have mental states do constitute logically adequate criteria for the presence of those mental states, then there is no sceptical problem (at least, the one the sceptic used cannot be generated).

Critical Examination of Strawson's Argument:

Premises 1. and 4. seem to me to be unproblematic and true. Furthermore, they are admitted to be true by the sceptic. These premises are therefore not the object of contention. Thus, (given that Premises 1. and 4. are true) it is Premises 2. and 3. which are the pivotal premises; if these are true, Strawson has, at least, dissolved (solved) the traditional problem of other persons. I shall then concentrate on Premises 2. and 3.

With respect to Premise 2. (Self-ascription meaningfully takes place only if other-ascription meaningfully takes place), first it is not the case that Strawson must be interpreted as relying upon some general form of Excluded Opposites Argument. If he were, his premise would thereby become problematic. For there is continuing controversy
concerning such type of argument. Secondly, we must keep in mind that Strawson is doing descriptive metaphysics. He is going to rely upon and use our actual conceptual scheme. So, he should not be interpreted as saying for example, "One logically (in the usual sense) cannot adopt some conceptual scheme wherein a use for (meaning for) the concept 'you' might be available without a use for a concept of 'others'." Rather, he should be interpreted as saying, in our conceptual scheme, "oneself" has a use (is meaningful) only because it has a meaningful (useful) contrast; namely, "others". Again, Strawson can be seen as presenting a conceptual law or truth.

It seems to me that Strawson is clearly correct here. Indeed, it is very obviously true that the concept of others (other persons) does have a use (uses) in our conceptual scheme and that this use is a use which arises partly from the practical need to refer to persons other than ourselves. For example, in the moral sphere, obviously there is a need to distinguish between those things, events, actions which are in the sphere of our responsibility, and those which are in the sphere of responsibility of someone else. Or, for example, in the ownership sphere, there is obviously need to distinguish between what I own and what is owned by other persons. And so on. One could provide endless examples of the need for and the occurrence of the distinction between oneself and others. In short, the use of "you" ("oneself") is "logically" dependent on the use of 'others'.

---

It is clear, then, that persons learn the use of the terms "I" and "you" and "me" and so on in the context of contrast to others. For example, my three-year-old daughter already distinguishes between what is "hers" and "mine" with what is that of "others" (Bill's, or Daddy's, or...). She learned the use of "hers" or "mine" or "her own" or "Paula's" only by contrast to what belongs to some other person or persons.

Now a critic might say, yes, but she need not have so learned the use of "mine" or "hers", and so on. She could have learned this use without any other persons having been involved, without any other persons having existed. So, the use of "mine" or "hers" is not necessarily bound up with any other persons or with any reference to others (any use of the concept of "other").

I have three objections to the above argument:

0.1. First, it is in fact false that my child could have learned the use of "her" or "oneself" or "me" even if no persons other than she existed. For if no person other than she had ever existed, she would not exist. That is, it is physically impossible that my child could have learned the use of words like "her", "me", and so on had no other person ever existed; for, surely, my child (as far as is now known concerning the natural laws of generation) could not exist unless she was born and she could not be born unless there was some person who conceived her (and thus existed!)

0.2. Secondly, she could not have learned the use of "her" or "mine" which she has without the contrast with others, because it is obvious that (it is a phenomenon that) the concept of "mine" we are
speaking of did develop as a term of contrast to others (persons other than herself) and thus if one developed some sort of concept of "oneself" without this contrast, it simply would be a different concept. That is, a part of the "logic" of the concept "you" ("oneself") or "me" and so on, is its use in differentiation among persons.

0.3. But, thirdly, even if one could somehow, by some incredible science fiction story, devise some way of showing how a concept of "oneself" could be meaningful without the existence of other persons or without the existence of the contrasting concept of "others", it is obvious that (it is a phenomenon that) we do develop our concept of "ourselves" in the context of contrast to things not ourself, which we call "other persons". So, once again, Strawson's premise 2. is merely an accurate recognition of this phenomenon.

But perhaps the sceptic could reply to these last two objections as follows: "While I grant that it is a phenomenon (he will surely not deny this) that the use of the terms "you", "I", and so on, do appear to have obviously developed in the context of the contrast to other persons, nonetheless, this is an illusory phenomenon."

Now the question becomes, does the sceptic have good reasons (or a good reason) for suggesting that this phenomenon is illusory?

He might suggest that these things we call "other persons" are merely robots. But, then, we ask him to produce evidence that they are. This he cannot do. For he cannot point to a robot that can duplicate all the things that can be done, said, and so on, by the sort of entity we call "other persons". But he persists that
these things we call "other persons" are merely robots. Then, we ask him what a robot is and his answer reveals that he is using the term robot to refer to exactly the sort of being we call an "other person". We can then merely suggest to him that if he wishes to call what we call "other persons", "robots", that is fine. But he has gained nothing. For we are left still with the view that "I" develops in the context of contrast to "other persons", which "other persons" the sceptic simply wishes to call "robots."

But now the sceptic insists that the matter is not linguistic. "What I want to say", says he, "is that you cannot know that those things you call "other persons" are other persons; that is, you cannot know if they are beings like yourself, with a mind as well as a body. You cannot know this because you cannot know if they have a mind."

If the sceptic argues thus, I merely point out to him, first, that he is using a Cartesian theory of what a person is (a mind and a body), which theory we have found to be very inadequate; secondly, I refer him to chapter viii, where it is shown that we do have empirical access to the mental characteristics (put linguistically, to P-predicates) of a person; and, thus, we can know if an existent entity is a person (a being such that both M-predicates and P-predicates appropriately and truly apply).

Premise 3. If it were true that behavioral criteria never could indicate meaningfully that some other person has any P-predicate, there would be no use for P-predicates with respect to any
other person. But, clearly, it does make sense (it is useful) to P-predicate with respect to other persons. That is, a look at our actual conceptual scheme shows that it is meaningful (useful) to P-predicate with respect to others. But what is a necessary condition of behavioral criteria sometimes meaningfully (usefully) indicating that (for example) "John is depressed" or "Jack is in pain"? Clearly, one necessary condition of this is that sometimes behavioral criteria do constitute logically adequate criteria for the correct ascription of P-predicates. Strawson is correct.

But could not the sceptic argue that Premise 3. is false, because the antecedent (B) is true, yet the consequent (C), false? That is, could he not suggest that while it is meaningful (useful) to ascribe mental characteristics to what are called "other persons" (to engage in other-ascription), yet there is still no good evidence whatever, in any case, that such ascription is correct? Just as it is useful to ascribe existence to God (for example, for keeping one's job, for getting charity money, for being well-received in certain circles, and so on), yet there is no good evidence whatever that such ascription is correct.

I shall argue that the sceptic cannot plausibly make the above objection. For, what the sceptic is doing is relying upon the phenomenon (a phenomenon the sceptic admits is non-illusory) that it is useful to ascribe states of consciousness to "others". But clearly he is here referring to the phenomenon that such ascription "works"; that is, often when one ascribes states of consciousness to
others, others do what one wants, or agree that such ascription is correct, or such other-ascription enables one to better understand and control these "others", and so on. But the question is, how could such other-ascription "work" time and again (as it does) if no such ascription is ever correct? Surely the most plausible explanation for why such ascription is often successful (in the sense of "working") is because such ascription is often successful (in the sense of "correct"). But if it is correct, then it is plausible to suppose that our ways of telling if a person has a particular mental characteristic (P-predicate) are sometimes logically adequate.

The alternative to holding the latter is to suppose that one is just incredibly lucky. That is, one would have to hold that time and again when one other-ascribes, though one never has good evidence to do so, one's other-ascriptions just happen to turn out to be successful (in the sense of "working"). The extreme implausibility of this alternative and the plausibility of holding that other-ascription is successful ("works") because it is often successful (correct), is sufficient, I believe, to refute the sceptic's objection. Thus, Premise 3. is true.

I have now argued, I believe successfully, that all of Strawson's premises are true. And clearly the form of the argument is valid. Thus, Strawson has produced a sound argument which shows that the sceptics' argument to show that our knowledge of other persons is illusory, cannot be generated. Thus, it has not been shown that our knowledge of the existence of other persons is illusory. We are, then, justified in treating our (apparent) knowledge as a fact.
Lastly (concerning PR.2.), I wish to try to provide a clearer understanding of the nature of P-predicates and thereby produce more evidence that other-ascription can plausibly be supported by empirical ways of telling sufficient to produce good evidence. This, again, will show how on ST the existence of other persons can be known.

First, let us take an example of a P-predicate; for example, "is depressed". Is the referent of the phrase "is depressed" an entirely private event or set of events? No. Says Strawson, "x's depression is something, one and the same thing, which is felt, but not observed, by x, and observed, but not felt, by others than x". Now, in light of this, self-ascription of depression can be evidential, ("I feel depressed; therefore, I am depressed"); and other-ascription of depression can be evidential (He is acting like he is depressed; He is speaking like he is depressed; He does not lie; and so on). In other words, according to Strawson, depression has both a public and a private dimension. It is something felt (private); and it is something observed (public).

Furthermore, the teaching (and learning) of the use (meaning) of the concept "depression" involves both its private and its public dimension. It is not the case that we somehow learn the use of depression privately and then somehow jump the gap between its private use (self-ascription) and its public use (other-ascription) by taking behavior as a sign of the private depression of another.

---

19*Individuals*, p. 105.
Not only does one not do this in terms of how one learns the use of the concept "depression"; Strawson's point is that one could not do this. For, one cannot go from one's own case ("my depression") to an other's depression, unless one already has a concept of an other. The logic of "oneself" ("my") and "others" is interdependent. But to have a concept of others is already to have bridged the gap. And Strawson's view of the P-predicate "depression" is recognition and explanation of how the gap is already bridged. Namely, there is no gap. The referent of the predicate is both felt and observed. Thus, one does have criteria of a logically adequate sort for the case of the depression (for example) of an other person. But it is precisely the assertion that behavioral criteria do not provide good evidence for the existence of the mental states of others which is used to justify the claim that we cannot know of the existence of other persons. I conclude that Strawson has solved (dissolved) The Problem of Other Persons.

PR.3. Before discussing specifically how Strawson deals with The Problem of Solipsism, two preliminary remarks are necessary. First, Strawson distinguishes between two types of Solipsism: 1) True Solipsism (the phrase is Strawson's) and 2) Philosophical Solipsism. The True Solipsist is a being with a solipsistic consciousness. That is, the True Solipsist is "one who simply has no use for the distinction between himself and what is not himself."20 It is not the case that the True Solipsist is one who believes that everything

20 Individuals, pp. 65, 66.
which exists is himself or a state of himself. The latter is the Philosophical Solipsist.21

Next, Strawson does not interpret the problem of solipsism in the traditional way (or, if you wish, in the way I have interpreted it). That is, he does not recognize the issue of whether only oneself exists and how it is known that something other than oneself exists as a "genuine issue at all".22 According to Strawson, he is going to "appropriate" the name ("the issue or problem of solipsism") for the following sorts of issues: 1) Each person does, in fact, distinguish between himself and what is not himself or a state of himself. In what way do persons make this distinction; and why do they make it this way?23

I am not here interested in this significant revision of the traditional way of viewing the problem of solipsism. I am interested in the traditional questions: 1) Can one know of the existence of something besides himself? and 2) How can one know this (if one can)? What could Strawson say about these issues?

In light of the arguments on PR.2, Strawson would be justified in observing that in so far as he has solved The Problem of Other Persons, he has solved The Problem of Solipsism. For, if his argument is successful (and I have argued at length that it is), then he has justified our belief that we know of the existence of other

---

21Ibid., p. 65.
22Ibid., p. 81.
23Ibid.
persons. Furthermore, he has there explained how one can know of the existence of these things other than ourself. Thus, he has thereby solved The Problem of Solipsism.

**PR.4.** Strawson makes only one reference to The Mind-Body Problem in *Individuals*. And this reference is merely a passing comment. He writes, "Earlier, when I was discussing the concept of a pure individual consciousness, I said that though it could not exist as a primary concept of a person (so that there is no mind-body problem, as traditionally conceived) yet it might have a logically secondary existence." 24

From this comment and Strawson's view on CT, on the primitive character of the concept of a person, and so on, I shall try to construct a Strawsonian viewpoint on this issue. After developing this, I shall argue that in contrast to PR.2. and PR.3., Strawson has not dissolved the mind-body problem as traditionally conceived; and that Strawson is left with a very similar problem (or set of problems).

What is the traditional problem? It has two major aspects. First, it is the problem of distinguishing between the mental and the physical; secondly, it is the problem of explaining what the relationship is between the two.

But why would Strawson suppose that these problems do not exist? (He claimed, you remember, that "there is no mind-body problem, as traditionally conceived".) I would guess that Strawson's reason

---

24 *Individuals*, p. 112.
is as follows: This problem arose in its modern form as a result of the philosophical viewpoint of Descartes. Descartes' theory (as interpreted by Strawson), you will remember, distinguished between two different types of substance; mental and physical. The major distinction between mental and physical substance was that the latter was extended, the former was not. But Strawson has argued that Descartes' theory of self which views a person or self as a mind and a body, is an incoherent, incorrect theory. A person is not a mind and a body, says Strawson; a person is one thing; an entity such that both mental predicates and physical predicates both appropriately apply. Thus, Strawson sees no need to distinguish between mental substance and physical substance. There is no such problem. Since a person is not "made up of" two different substances, there is no problem of distinguishing two different substances. And, likewise, since there are not two different substances, there is no problem of discovering or explaining the relation between them.

My response to this reason (or argument) is as follows: I agree with Strawson that given the abandonment of the Cartesian dualism, there is no longer a need to distinguish mental substance from physical substance. But, it seems to me that once one then adopts a monism such as Strawson's, along with the admission that there are still two aspects to a person as reflected in the predication of two distinct types of predicates, M-predicates and P-predicates, one still has the problem of distinguishing between those two aspects of a person; and of explaining the relation between them. And
Strawson does not even try to deal with these problems. At most he simply makes a distinction between M-predicates and P-predicates, such that the appropriate predication of the latter imply the possession of consciousness on the part of that to which they are predicated; and the former do not. But, this is not all satisfactory, for, what distinguishes consciousness from the physical is still problematic. And, Strawson makes no attempt at all to make it less so.

PR.5. Again, Strawson does not go into detail about his solution to The Problem of Personal Identity. He does, however, make some remarks concerning this problem. What I shall do in this section is first, exposit his major comments which are relevant to this problem. Next, I shall pull these together and as best I can, give a plausible Strawsonian answer to this problem. That is, I shall argue that Strawson can solve the problem.

First, then, Strawson makes the following major comments relevant to this problem:

1. Once it is understood that persons are entities which perceptibly occupy space and time and thus can be identified and individuated in largely the same way as "other items having a material place in the spatio-temporal framework", the problem of personal identity is relatively easily solved.

2. The fact that a person has a physical dimension is what enables us to have empirical criteria for personal identity.

3. But the criteria for re-identifying persons is not the same as the criteria for re-identifying material bodies. (Recall that for Strawson, a person is not just a material body.)
4. There are ordinary criteria for personal identity. That is, persons do have criteria which they use, on the whole, unproblematically, to re-identify persons.

5. There is not just one criterion of personal identity. There are many.

6. In certain extraordinary circumstances, we might have to decide how to adjust our concepts of identity and re-identity.25

Now I shall try to pull the above together and give a direct Strawsonian solution to The Problem of Personal Identity. First, in so far as the problem is seen as the problem of explaining how we can decide that a given person is the same person, Strawson's answer is that we use a number of ordinary empirical criteria. For example, we use a visual criterion (does he look like John?). We use an auditory criterion (does he talk like John?). We perhaps sometimes even use an olfactory criterion, or a gustatory criterion, or a kinesthetic criterion, though much less frequently. And most often we use more than one of these. We, thus, use multiple criteria. But we do not use only these bodily criteria. We also use what might be called mental criteria. For example, we use data resulting from such questions as: "does he think like John?"; "does he get depressed like John does?"; "are his intentions those of John?"; and so on. So the criteria for re-identification of persons are not the same as the criteria for re-identifying ordinary material bodies. They include such criteria but do not exhaust such criteria. In short, the criteria for persons also involve what is unique to persons (as

25The above list is paraphrased from pp. 131, 132 of Individuals.
opposed to material bodies); namely, P-predication (the mental aspects of a person).

Now Strawson's theory escapes the problems generated, for example, by the Cartesian theory, because, it does not have the difficulty of re-identifying an immaterial entity before being able to re-identify a person. And it escapes the problem of the No-Ownership theory because it can correctly grant that re-identification of a person is not merely re-identification of a material body (that is, without reference to P-predication).

Furthermore, Strawson could observe that the concept of a person is the concept of an entity which is said to remain (one) through change. As Terence Penelhum has observed (with respect to Hume and this problem of unity):

There is another, closely related, mistake which Hume made. This is the mistake of thinking that for anything to be entitled to be called "the same" it has to remain unchanged from one period to the next. This is a muddle of two things that he himself distinguishes at one point, viz., the two distinct senses of the word "identical" or "the same". These are the numerical and the specific senses, as he calls them. Two things can be the same as one another in the specific sense, i.e., exactly alike in some respect, yet they will still be two things; but if they are said to be the same in the numerical sense they are being said to be not two things but one after all. These two senses are distinct from one another. Now to remain unchanged is to remain the same in the specific sense, i.e., to be now exactly as one was at an earlier time. But I can remain the same in the numerical sense without doing so in the specific sense--I can be numerically the same but changed. In fact, I cannot be said to have changed unless I am the same in the numerical sense....In the case of most things, the words we use to talk about them are words the meanings of which allow us or require us to continue to use them throughout certain changes, though not of course any changes. What kind of
changes can occur without our having to say that the
thing has ceased to exist and given place to something
else depends on what kind of thing we are talking about.
To know what such changes are is part of what it is to
know the meaning of the class term for that sort of
object. A house, or a person, is something which admits
of many changes before we would say it had ceased to
exist. To know what these changes are is to know, in
part at least, what the words "house" and "person" mean.

There is, then, no mystery to the unity or sameness of a per-
son. Persons are simply the sort of thing we say can remain one,
even though they change. (This is part of our conceptual scheme.)
And, to know just how much and what sort of change can occur to the
sort of thing we call a person before we say he is not the same per-
son is simply a part of what it is to know what a person is (that is,
to be able to correctly wield the concept of a person). For example,
we do not say of a person who has changed his hair color that he is
a different person. We say, "He is the same person, but he dyed his
hair." On the other hand, we do sometimes say of a person who has
radically altered in appearance and behavior, that "He is a different
person.", or "He is not the same person." In the latter case, we
might say, "In a way he is the same person (perhaps, in the numerical
sense), yet in another way, he is not the same (in the specific sense)."
And, sometimes, we just do not know what to say. And, thus, we might
simply have to make a decision as to if and how far we shall extend
the applications of the concepts "same" and "different".

Strawson can, then, solve PR.5.

---

CHAPTER X

STRAWSON'S THEORY AND GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES

As far as Strawson’s Person-Theory and III. is concerned, only P.1. (The Simplicity Principle) needs to be argued. P.2. (The Consistency Principle) once more remains idle. And P.3. (The Ad Hoc Principle) is not violated by ST.

Concerning P.1. then, does Strawson produce a theory which is simpler than the other theories which we have examined? I would like to begin by distinguishing two types of simplicity. I shall call the first type linguistic or conceptual simplicity; and the second type, ontological simplicity. Linguistic simplicity involves the nature of the concept or concepts involved in the theory. Specifically, it a theory of persons analyses the concept of a person as involving two other concepts (as does Descartes' theory), then a theory which analyses the concept of a person as involving only itself (as a basic or primitive—non-analysable concept) is conceptually or linguistically simpler. Strawson's theory is thus, in comparison with Descartes' theory, conceptually simpler. It is also conceptually as simple as a theory can be.

On the other hand, ontological simplicity is a function of the number of distinct ontological kinds or types recognized by the theory.

151
And, at first glance, it would seem that Strawson has produced the simplest theory of all; since, according to Strawson's theory, a person is one thing. Surely, no theory could be—in terms of number of ontological entities—simpler. And, as a matter of fact, CT and HT both claim that a person is a dual entity. Descartes' theory does so in the most obvious way by stating that a person is a mind (mental substance) and a body (physical substance). Hume does so less obviously. For, a Humian person is a set of perceptions in relation. Thus, unless Hume can show that relations are also to be identified with perceptions, which Hume does not do, his person is a dual entity as well. The No-Ownership theory, however, is monistic. A person is one thing; namely, a body.

While I do, then, agree that Strawson's theory is prima facie as simple a theory as can be desired (from the standpoint of ontological simplicity) nonetheless I wish now to argue that, in fact, Strawson's theory suffers from a lack of ontological clarity. After explicating this lack of clarity, I will attempt to remedy it. On one level, one can grant that a person is one thing to which both M-predicates and P-predicates appropriately apply; nonetheless, when the ontological question is pressed (What is a person? Of what does a person consist?) Strawson's answer is not that a person is a body; nor is it that a person is a mind; nor is it that a person is a body and a mind. It is none of these. But, then, what is it?

Strawson's response at this point would be to reiterate that a person just is that sort of entity such that M-predicates and P-predicates necessarily do appropriately apply. That is, the defining
characteristics of the type of entity designated "person" is that it is an entity to which necessarily both M-predicates and P-predicates appropriately apply. I believe this Strawsonian response is partially ontologically adequate, but only if one specifies further which P-predicates are necessarily appropriately applied. (This Strawson does not do.) By this I mean, it simply is not the case that an appropriate specification (definition) of a person is an entity to which (necessarily) both M-predicates and P-predicates apply. For, this would be true of, for example, a dog. Strawson gives "is going for a walk" and "is in pain" as examples of P-predicates. Now both these are appropriately predicated of dogs. So also are M-predicates, such as "weighs 10 stone" or "is in the drawing-room". Now it is true that there are some P-predicates which would not be appropriately applied to dogs; such as "believes in God".

Thus, if Strawson wishes to clearly distinguish persons from non-persons, if he wishes to suggest that a person is neither a body nor a mind nor a body and a mind but rather is a unique ontological type—unique ontologically distinguishable from not only trees and tables (for P-predicates are not appropriately applied to these) and from dogs and cats and such—then he must make a further distinction between P-predicates which are and P-predicates which are not uniquely applicable to persons. I shall call those P-predicates which are not unique to persons simply "P-predicates"; and those uniquely applicable to persons "P\_p-predicates". Some examples of "P\_p-predicates" would be as follows:
1. "believes in God"
2. "hopes Nixon will lose"
3. "worries about his college grades"
4. "remembers his third grade teacher"
5. "thinks about the Vietnam War"
6. "decides to go to Harvard"

With this revision, then, a person would be that unique type of entity such that necessarily both M-predicates are properly or appropriately applicable and P_p-predicates are properly applicable. This would distinguish persons from all other ontological types. And, so far, Strawson still retains an ontological monism.

However, there is another area of possible ontological unclarity. Does not the fact that a person is an entity such that two different types of predicates properly apply to it imply that a person has, as it were, two aspects or dimensions; namely, a physical dimension and a mental dimension? Now it is true that Strawson does not use either of the above terms ("aspects" or "dimensions"). And I am not suggesting that he must. I am merely raising the question. And I am raising the question because it seems obvious that if an entity is predicated of in two quite distinct ways, by two quite distinct types of predicate, then one would be justified in suggesting that that entity has two aspects or dimensions. For example, if a college is such that a certain set of predicates properly apply to one "part" of the college (A) (predicates such as, "has a traditional curriculum", "has one-hour classes", "uses books", "has a regular faculty" and so on) but not to another part (B); and a certain different set of predicates properly apply to another "part" of the college (B) (such predicates as, "has no curriculum", "has no classes", "has no regular faculty", "uses no books", and so on) but not to (A), then one would surely be
justified in saying that this is a college with two very different aspects or dimensions or "parts", (A) and (B). Now if this one college was defined as (A) and (B), then we have an example similar to the case of Strawson's person.

Now what I wish to argue is that Strawson can grant that a person has two aspects or dimensions as reflected in the two different types of predicates applicable to persons, yet he can still justifiably say that ontologically a person is one thing. Just as we could say, plausibly, that the college spoken of previously is one college. There is, nonetheless, a difference between the Strawsonian person and the college example. For (A) or (B) of the college could be identified independently. That is, if (A) or (B) somehow literally went out of existence, one nonetheless could make correct identifying references to the other; in short, one could identify the other. But in the case of the Strawsonian person, the relationship between the aspects is much closer, as far as identification is concerned. For, if the material aspect of a person went out of existence, one no longer could identify the mental aspect. For, identification of this does depend on identification of the person; which person would no longer exist, if the material aspect no longer existed.

On the other hand, the college example is relevantly similar in two respects. First, if (A) or (B) of the college no longer existed, the college (which was defined as (A) and (B)) would (logically and actually) no longer exist. Just as if the material aspect or the mental aspect of a person no longer existed, the person would (logically and actually) no longer exist. For, the material aspect
of a person is merely that aspect which is referred to by way of M-predication. And the mental aspect of a person is merely that aspect which is referred to by way of P-predication. And for Strawson a person is an entity such that both M-predication and P-predication necessarily properly apply.

While one can thus grant Strawson a case both for his theory of self being linguistically the simplest and in a sense, ontologically the simplest, nonetheless, until Strawson speaks more fully, if he will, on the ontological question of the nature of the mental aspects of a person and physical aspects of a person, his theory will remain problematic in respect to at least one aspect of its simplicity. For, how much of a gain, in terms of ontological simplicity, has he made over, for example, Descartes' theory. For, Descartes could also claim that a person is "one" thing, namely, a body (A) and a mind (B); and (A) and (B) are merely different "aspects" of this one thing, a person.

Yet, I believe Strawson is, at this point, in rather a dilemma. If he shows that there is in fact a significant difference between those aspects of a person referred to by P-predicates and those referred to by M-predicates, then his theory would be a form of dualism and thereby be no simpler, ontologically, than Descartes' theory. If, on the other hand, he suggests that there is no significant difference "gotten at" by the differences in P-predicates and M-predicates, then while he retains an ontological simplicity, he does so by generating a linguistic mystery. For, now the question arises, why distinguish between M-predicates and P-predicates? That is, if the distinction
between M-predicates and P-predicates is not accounted for by a difference in the reference or applicability of the predicates (such as P-predicates refer only to or are applicable only to entities which possess consciousness) then (a.) Strawson has thereby abandoned the basis of his distinction (for it is that P-predicates refer only to or are applicable only to entities possessing consciousness) and (b.) he is left with a mystery; namely, why there is a distinction made between M-predicates and P-predicates.

In summary, ST is conceptually as simple a theory as is possible. Furthermore, ST is ontologically simplest in that a person as a type of entity is unique. That is, a person is singled out as a unique ontological type via M- and P-predication. But, ST does suffer from a possible inherent dualism as a result of the distinction between the types of predicates applicable to a person. If they are type-different, this suggests that a person consists of two different types of entity (material and conscious?). Yet, if they are not type-different, how are they different? Or, if they are not different at all, why does Strawson treat them as different; and, indeed, as type-different? ST, thus, both succeeds and fails (though of course, in different respects) in relation to $P$.3.
CHAPTER XI

CONCLUSION

In this last chapter I wish to do three things: 1. make a rather detailed comparative evaluation of the theories of self which I have examined; 2. make some suggestions as to how Strawson's theory could solve its remaining problems; and lastly, 3. summarize the value but also point out the limitations of this study in the philosophy of self. I begin by making a detailed sketch of the comparative merits of Strawson's theory.

With respect to I. (PHENOMENA ACCOMMODATION), ST is almost completely successful. Its major advantage is its ability to grant what is obviously correct, namely, that there are clear, ordinary, and generally successful ways of re-identifying, identifying, and individuating persons. The only other theory that Strawson examines that is able to accommodate these phenomena (PH.2., 3., and 4.) is NT. Both CT and HT fail with respect to all three. ST is also able to accommodate our ordinary language uses of the term "person". In some respects, none of the other three could do so.

And lastly, ST could accommodate PH.1. but only partially. Neither CT nor HT nor NT could deal with PH.1. in a completely satisfactory way. Clearly, then, ST is the most successful of the four
theories as far as PHENOMENA ACCOMMODATION is concerned. ST could accommodate four of the five Phenomena; and the fifth it could partially accommodate. CT and HT could satisfactorily accommodate none of the five. NT could accommodate only three of the five.

Turning to II. (PROBLEMS TO BE SOLVED), ST is again comparatively most successful. All the theories are successful with PR.1. (The Problem of the Nature of a Self). But only ST is able to solve (dissolve) The Problem of Other Persons (PR.2.) and The Problem of Solipsism (PR.3.). These solutions I would consider to be among the most significant aspects of the success of Strawson's philosophy of self. The Cartesian theory generates both problems and is unable to solve them. And HT and NT cannot escape them. On the other hand, none of the theories is able to solve The Mind-Body Problem (PR.4.). Only NT and ST were successful at solving The Problem of Personal Identity.

ST, then, is the most successful of the theories with respect to II. (PROBLEM SOLVING). CT failed to solve four of the five problems and had two further unresolved objections. HT failed to solve four of the five problems and had two further unresolved objections. NT failed to solve three of the five problems and had two additional unresolved objections. ST failed to solve only one of the five problems and had no further unresolved objections.

Lastly, in respect to III. (GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES) as far as P.2. (The Consistency Principle) is concerned, all the theories are equal; for P.2. is, as of now, idle. There are
no well-established theories (scientific or otherwise) directly relevant to the philosophy of self. Only CT had recourse to Ad Hoc explanations (P.3.). And only NT had as simple a theory as ST. So, all in all, ST and NT are equally more successful than CT and HT with respect to III. (GENERAL EXTRINSIC THEORY PREFERENCE PRINCIPLES).

It is clear, then, that as judged by the Theory Preference Structure which I developed and defended in chapter 1, Strawson's theory is the most plausible of the theories examined, the most defensible of the theories examined, the preferable theory among those examined.

I wish now to briefly explore the question, "How can Strawson's theory be made even stronger?" How can it remedy its remaining defects? An obvious starting point is to ask what, in sum, are its remaining defects? They are: 1. While ST could accommodate PH.1. from a linguistic point of view, it could not from an ontological point of view; for it did not satisfactorily distinguish between the mental and the physical and thus could not but leave the actual mechanism of agency, which was said to involve intentions (a mental dimension), in mystery. Until this is clarified, it is not able to be decided if a Strawsonian person can do things. 2. ST could not solve or dissolve PR.4. (The Mind-Body Problem). This defect of Strawson's theory is closely related to the first defect. For, it is once again the failure of Strawson to satisfactorily answer such an obvious and pressing question as, "Why is there a linguistic distinction between M-predicates and P-predicates? What is the ontological reason for such a distinction? What is it about a person
which generates such a distinction?" To answer that $P$-predicates imply the possession of consciousness on the part of that to which they are ascribed, merely re-generates the question in the form, "What is it which distinguishes consciousness from the physical (or from non-consciousness)?" The further question, "What is the relationship between the consciousness of a person and the non-conscious?" is, thus, also left unanswered. 3. Lastly, and once again closely relatedly (to defects 1. and 2.), while Strawson's theory is conceptually as simple as possible ($P.1.$); and, while, in one sense, Strawson's theory is also ontologically simple, nonetheless, because of the lack of clarity with respect to the mind-body problem Strawson's theory suffers from a possible latent dualism. Until this dimension of Strawson's theory is clarified, ST is not as acceptable or successful as it could be.

If, then, I am correct in my judgment of the success of ST in relation to $PH.2.$, $3.$, $4.$, and $5$; $PR.1.$, $2.$, $3.$, and $5$; and $P.2.$ and $3.$, Strawson's theory has only one major defect, which defect affects $PH.1.$, $PR.4.$, and $P.1.$ That defect is the lack of clarity which surrounds Strawson's treatment of (or rather, lack of treatment of) what I call the two "aspects" of a person: the mental (conscious; states of consciousness) and the physical. For, once this is clarified, answers could be forthcoming as to 1. Can a Strawsonian person do things? ($PH.1.$) 2. What is the distinction between the mental aspects of a person and the physical aspects? And what is the relation between them? And 3. Is Strawson's theory in one sense, at least, on ontological dualism?
I now wish to suggest that ST could solve this problem by a defense of the Identity Theory. I, of course, cannot undertake a defense of this very controversial theory in this study. Indeed, such a defense would require at least one more full dissertation. I can, however, show how such a defense would remedy the very significant remaining defect in ST.

One major version of the Identity Theory holds that while the logic of statements using mental concepts or terms is different from the logic of statements not using such terms (using terms involving physical concepts), nonetheless, as a matter of contingent fact, the referent of both types of statement is the same; namely, brain processes. That is, it is granted that true and meaningful statements such as, "I am now thinking" or "He desires that you do it", cannot be translated into statements which are still true and have the same meaning, yet do not involve mental terms or which involve only physical terms. But, while this is granted, it is denied that there is something unique, something uniquely psychic, referred to by these statements. Rather, it is asserted that things like thinking, hoping, calculating, sensing, and so on do exist, but are, in fact, identical with physical (brain) processes.¹

The question now is, how could Strawson's theory gain from an adequate defense of the Identity Theory? First, in terms of the agency objection to ST, if the Identity Theory were correct, then there simply

is no mystery concerning "doing". For, "doing", like all else, is a physical act. It may well involve intentions. But, according to the Identity Theory, intentions are physical (brain) processes. Thus, there is no ontological mystery concerning doing; doing is simply a physical act of a certain sort.

Secondly, with respect to the Mind-Body Problem, if the Identity Theory were correct, there is no Mind-Body Problem at all; for, there is no Mind. And, the ontological mystery of the two aspects of a Strawsonian person is dispelled. For, ontologically there are not two aspects, there is just the physical aspect; for, the so-called mental aspect (consciousness) is, in fact, a physical (brain) process. Yet, Strawson can retain the distinction between M-predicates and P-predicates. For, P-predicates can still be those predicates which imply the possession of consciousness on the part of that to which they are ascribed. That is, it is still true that only certain types of entities in the world are, in fact, conscious (have this physical property). The only thing that the Identity Theory adds is an analysis of consciousness which reveals it as physical. And the justification for distinguishing M-predicates from P-predicates can simply be the fact that there is, in fact, a distinction between entities which are conscious and those which are not; and more specifically, the justification for distinguishing \( P_p \)-predicates from P-predicates and M-predicates can be that \( P_p \)-predicates distinguish even more specifically a type of entity; namely, persons.
Thirdly, and lastly, a successful defense of the Identity Theory can solve the problem of the possible latent dualism of ST. For, if the Identity Theory were correct, there are not two aspects to a person at all (except in so far there are two distinguishable types of predicate, P- and M-predicates). A person is, ontologically speaking, one thing; namely, a physical entity.

I wish, in concluding, to point out what I believe to be the values yet also the limitations of this extended study in the philosophy of self. In terms of values, there are many: 1. I have herein developed and defended a set of criteria for rationally judging the comparative success or failure and degree thereof, of any theory of self. 2. I have, in the context of this set of criteria, critically examined four major theories of self (CT, HT, NT, and ST). 3. As a result of 2., I have produced a good (comparative) case for claiming that Strawson’s theory of self is correct (that is, is the best theory of self now available). In other words, I have created a justification for the claim that Strawson's theory of self solves what I would call the philosophical problem of the self. 4. As a side result of this study, I have produced some evidence that some philosophical problems can be solved by careful attention to and careful description of ordinary language and our actual conceptual scheme. 5. On the other hand, I have argued that the problem of self, itself, cannot be simply solved merely by such conceptual description. Rather, it must be solved by the use of phenomena as well as linguistic data; and by being judged in the context of a rationally defensible set of criteria developed for the judging of theories of the self. 6. I have
produced some evidence that progress occurs in philosophy (as it does in science).

While I believe I have done a number of valuable things in this study, there are some limitations which I do recognize and which I hope the reader recognizes. First, for example, I have not examined all of the major theories of self which have been developed. Thus, it is possible that one of those other theories is as successful as Strawson’s theory, or even more successful. On the other hand, I have examined four of the most important and influential of the actual theories of self that have been developed.

Secondly, while I have tried to be as complete as possible with respect to the criteria I developed in chapter i, I may have excluded certain criteria which would be relevant to the judging of a theory of self. Though I, of course, do not believe I have been incomplete, if I have, this may affect the decision as to the best theory.

Thirdly, I could, of course, be mistaken with respect to some of my arguments: defenses, objections, and/or counter-objections. If so, this, too, could affect the outcome.

I have pointed out the limitations as well as the values of this study because I wish to re-affirm the conception of philosophy which has guided this study. One of the major tasks of philosophy is to solve philosophical problems. The solution to those problems is generally in the form of a theory. And the theory is to be judged in the light of the PHENOMENA TO BE ACCOMMODATED, THE PROBLEMS TO BE SOLVED (and the Objections to be resolved), and, THE EXTRINSIC THEORY
PREFERENCE PRINCIPLES to be conformed to. The decision as to the correct (best) theory with which to solve a philosophical problem is, thus, always a tentative decision (as is true of theories in science). For, the decision is a function of the Phenomena, Problems, Objections, and Preference Principles. And if new Phenomena become available, or if the new Problems arise, or if new Objections are generated or if new Phenomena affect the applicability of the Preference Principles, then one's decision as to the best theory may be affected.

With the above in mind, I conclude that P. F. Strawson has produced a (comparatively) successful theory of self and has thereby solved the philosophical problem of the self. The extended defense of that claim, you have just examined.
SELECTED BIBLIOGRAPHY

Books


Articles


Bradley, M. C. "Mr. Strawson and Scepticism." Analysis, 20 (October, 1959), 14-19.


