



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

P E R S O N A L I D E N T I T Y

( A Critical Survey of the Problem from  
John Locke to the Present Day )

By

P.K.Mohapatra, M.A.

A thesis submitted in partial fulfilment of the requirements  
for the Degree of Doctor of Philosophy at the Univ. of Keele.

April 1979

Dr, P, K, MOHAPATRA

This thesis is published in a slightly reduced version  
under the title "Personal Identity" .

Printed in India at the Utkal Rashtrabhasha Prachar  
Co-op. Press & Publishing Society Ltd.,  
Rashtrabhasha Road,  
Cuttack - 753 001.

1983.

"...if things were quite different from what they actually are - if there were for instance no characteristic expression of pain, of fear, of joy; if rule became exception and exception rule; or if both became phenomena of roughly equal frequency - this would make our normal language-games lose their point."

(Philosophical Investigations, I §142)

# CONTENTS

Preface	page (i)
An Abstract	page (iii)

## Part - I

### Chapter 1 INTRODUCTION

Section (i)	<u>The Problem</u>	page 1
Section (ii)	<u>A Special Problem</u>	page 27

## Part - II

### Chapter 2 THE TRADITIONAL APPROACH

Prologue:		page 55
Section (i)	<u>Locke and Hume: the memory theorists</u>	page 57
Section (ii)	<u>Butler and Reid: the intuitionists</u>	page 70
Section (iii)	<u>The Views of Kant and William James</u>	page 81

### Chapter 3 THE CONTEMPORARY APPROACH

Prologue:	<u>A Problem of Criteria</u>	page 97
Section (i)	<u>Criterion</u>	page 99
Section (ii)	<u>Bernard Williams: Bodily Continuity</u>	page 112
Section (iii)	<u>Shoemaker and Wiggins: A preference</u> and a qualification	page 140
Section (iv)	<u>Derek Parfit: memory with a</u> difference(?)	page 170

## Part - III

### Chapter 4 THE TWO CRITERIA

Prologue:		page 187
Section (i)	<u>Bodily Continuity Re-examined</u>	page 189
Section (ii)	<u>Memory: Not an independent criterion</u>	page 202
Section (iii)	<u>Disembodied Persons</u>	page 218
Section (iv)	<u>More About Disembodied Persons</u>	page 237

### Chapter 5 CONCLUDING ESTIMATE

Section (i)	<u>Conflict of Criteria and the</u> <u>Relevance of Puzzle Cases</u>	page 255
Section (ii)	<u>Must There Be An Answer ? - The</u> <u>Importance of Personal Identity</u>	page 271

<u>Appendix 1</u>	page 288
<u>Appendix 2</u>	page 291
<u>Notes to the text</u>	page 296

P R E F A C E

In the planning of this thesis I have, with apologies, slightly deviated from normal practice. The numbered notes containing references to relevant literature have been supplied together at the end of the text; but certain other notes, with asterisks, are given in the text at the foot of the page where they occur. This has been done in view of the fact that these latter notes are more important to the body of the text and need prompt attention.

Among the many philosophers whose writings have influenced me and helped me to formulate the arguments of this essay are particularly Shoemaker, Wiggins and Williams. Despite my qualified disagreements with their conclusions, their original and stimulating works have supplied much of the elements of my own view that is going to be presented.

One single philosopher whose writings on the topic have helped me to see the present problem in its right light is Professor R.G.Swinburne. Besides this, I am grateful to him for his constant, hard-working and helpful supervision of my work, without which the present work could not have taken shape. I am, however, entirely responsible for whatever errors and inadequacies are contained in this work. I am also grateful to Dr.Brian Smart through his writings and also through direct discussions with him on the topic. Dr.Smart and Prof.J.L.Mackie deserve my special gratitude for their valuable suggestions after examining the earlier version of the thesis.

I am obliged to the Commonwealth Scholarships Commission in the U.K. for the generous grant to enable me to go and work at Keele, and also

to the authorities of the Utkal University ( India) for granting me the necessary leave. My thanks are also due to Mrs Kathleen Hampton for typing the first draft of this thesis and to Mr. Varghese Marzelin for typing the second draft.

Utkal University,  
Bhubaneswar (India)

Prafulla K.Mohapatra

April 1979

AN ABSTRACT

The problem of personal identity, as I propose to discuss herein, is the problem of what, if anything, makes a person the same person at different times. It is a problem because although many things about a person go on changing - his physical features as well as his psychological features - through time, this normally does not affect our saying that a later person is the same person as an earlier person (This talk of 'a later person' being the same person as 'an earlier person' in terms of which the problem is raised and discussed, as it usually is, need not be contradictory or otherwise pleonastic; for they are only a shorthand device for the expressions 'a person picked out at a later time' and 'a person picked out at an earlier time'.) Is there, then, something unchanging about a person which is the bearer of his identity through time? Various answers have been suggested to this question from John Locke's time up to date, though hardly any answer has been satisfactory. In the present work, I want to consider why there should be a problem - which, to many, is a very special problem - about the identity of persons if none so special seems to be there about that of most other things, and to critically assess the various answers suggested. I will also try to find out what made different philosophers give different - often conflicting - answers to the problem, and thereby, to point out why these answers were unsatisfactory. By way of such critical assessment, there will emerge my answer to the problem which I will claim to be free from the difficulties and limitations inherent in the ones I examine.

Incidentally there is another form in which the problem can be raised and this is the problem of synchronous identity. This involves the interesting issue of whether two (or more) contemporaneously

identifiable selves are one and the same self. I shall not here be concerned with this aspect of the problem since it forms a separate problem from that of re-identity of persons which is chosen as the subject of my thesis.

In the interest of clarity and convenience chronology may not be strictly adhered to - though every attempt will be made to stick to it as far as possible.

The thesis will divide into three parts dealing, roughly, with the nature and source of the problem, how it has been looked at and dealt with by different philosophers and what a proper analysis of the problem will amount to. These three parts, in turn, will spread over five chapters.

The first chapter will be an introduction in which the problem will be stated and the nature and source of it will be clearly brought out. The peculiarity of the problem of personal identity, it will be pointed out, is due to a more intimate connection between the concept of a person and the criteria for the identity of persons, and also to the fact that persons are self-knowers. It will then be maintained that an approach to the ~~par~~ problem will be on the right line if it is taken as a problem of specifying the criteria for making personal identity judgments - and not as an attempt to define personal identity since, it will be argued, no satisfactory non-trivial definition of the latter is possible.

I shall then go on to consider, in the next two chapters, the way the problem presented itself to the traditional philosophers and the way it has been looked at in recent writings. Some possible connections and distinctions between the different views will be uncovered. The detailed scheme of these two chapters give the philosophers whose

views are being considered. I will not pretend that the list is in any way exhaustive, but I do hope that it presents the major links in the chain. Other writers and commentators will be given due attention in the course of the text. One major distinction between the traditional approach and the contemporary approach will be brought out in the following way: whereas in the former the problem was looked at (so I will argue) pre-eminently as one of definition, in the latter the question generally has turned on the problem of specifying the criteria to be used in making personal identity judgments. This difference in approach will be shown to explain the relative clarity of the contemporary literature and the somewhat vague and even paradoxical nature of most traditional answers.

After having thus shown that our problem is one of criteria, I shall set myself to the natural task of reviewing the status of the two main criteria of personal identity, namely similarity of memory claims (with or without that of personality and character) and bodily continuity (which includes spatio-temporal continuity). This will be my concern in the 4th chapter. Attempt will be made to show that bodily ~~iswarity~~ identity is an independent and the primary criterion of personal identity. But bodily identity, I shall contend, should not be taken in so rigid a sense as Williams, for example, has taken it but that it should be qualified to take the spatio-temporal continuity of whatever may be the physical basis of what I shall call 'the personal faculties'. Such lines have recently been suggested, notably by Wiggins, Shoemaker and Parfit, in some form or other, by imagining the possibility of brain (and/or split brain) transplants. But I will argue that cases of brain transplants show, not that memory continuity or "psychological continuity" - to the exclusion of bodily identity - is the criterion

of personal identity as these philosophers seem to think, but that such possibilities can be so interpreted as to preserve bodily continuity as the necessary condition of personal identity. On the other side, arguments will be given to substantiate the well known ~~ex~~ claim that memory cannot be an independent criterion of personal identity and that it has to depend, for its successful application, on the bodily identity criterion. This will oblige me to examine the alleged possibility of disembodied existence, for if such existence is possible for persons this would supply a good reason for memory being an independent criterion. My attempt will be to argue against such possibility. My argument will not be designed to show that this idea is logically incoherent or straightforwardly nonsense, but to ~~h~~ show that it is unreasonable and, more particularly, that it does not show, what it purports to show, that memory is the sole (or even the primary) criterion of personal identity.

In the last chapter, I will consider some cases where the two criteria are said to conflict and where, consequently, there seems to be no right answer to the problem; judging the extent of the bearing of these "puzzle cases" on personal identity I shall argue that these cases create no conceptual problem and so justify no plea for revising our present concepts of a person and personal identity. I shall then state the importance of personal identity and discuss critically why it is that we demand all-or-nothing answers to personal identity questions and why it is that such answers are not possible in some cases.

PART - I

## Chapter 1

### INTRODUCTION

#### Section (1): The Problem

Problems of philosophy usually arise out of attempts to explain some facts of common experience. Phenomena that are otherwise taken for granted appear to be bewildering and paradoxical when a rationale for them is asked for. It appears as if we have no right to say what we do say about them. Things that so commonly pass as brute facts often become matters of interpretation, depending on how one wants to describe them. And how one wants to describe a phenomenon is quite often guided - consciously or unconsciously - by the basic philosophical assumption or assumptions of the philosopher who describes them. But a proper method of philosophising should consist in explaining the phenomena without any preconceived notions whatsoever and yet in keeping the explanations as near to the facts as possible. Any basic philosophical assumption, far from being wrong or incorrect, may be the right theory and may have great explanatory value; but the philosopher's task is to show that it is so - rather than assume it and explain the phenomena by its means. My purpose in the present work is to pose the problem of personal identity, expose the motives of the different philosophers underlying their answers to the problem and to propose what I hope to be a proper philosophical explanation of the problem in the sense just indicated. In the present part (i.e. in the first chapter) I shall confine myself mainly to the first of these tasks and defer the two others primarily to the two subsequent parts of the proposed thesis.

It is a very common experience with all of us that we recognise our relatives, friends and acquaintances. We also are commonly able to say whether or not a particular person is the (earlier) person he says he is. This process of recognition or reidentification of persons involves a

lapse of time during which some changes - often a great deal - might, and do, come over the persons. The changes relate to the physical as well as the psychological features of persons, for example: my memory, character and personality have changed considerably. I now remember a lot more than I did at any earlier time, I have also forgotten many things which I did remember earlier; my character too has undergone some changes - my way of approach towards things and people have changed, some of these changes are quite considerable and some not; as regards my personality, I have been adding something to this possibly with every increase of experience and learning - in some sense, the changes in my personality have been remarkable (evident from the contrast between the rustic manners as a school boy and the somewhat refined ways of an urbanized youth). On the physical side, my body has obviously grown bigger and stronger, my complexion and appearance have changed considerably; and if the theory that in every seven-years period (or so) all the molecules of a human body are replaced by different ones is true, then there is not a single particle of matter now in my body which was there in (or is common to) the body with which I touched ground or even the body which I am said to have had ten years back. These are obvious things that have happened to me and do happen to all persons. Yet we do say, and believe, that persons remain the same over periods of time. Reflecting this, a recent writer describes the problem of personal identity as the problem of trying "to justify a practice which seems at first sight to be strange, and even paradoxical. This is the practice of talking about people as single beings in spite of the fact that they are constantly changing, and over a period of time may have changed completely".<sup>1</sup> As I have said, the "practice" referred to is very common indeed; it is not merely a "practice of talking" but a practice that carries with it a strong conviction about the nature of persons and personal identity. The conviction is that persons are single beings and that a

person continues to be the same person throughout (what we call) his life time. We can easily notice that this way of putting the problem involves two distinct issues. One concerns the nature of persons - whether persons are, as they are commonly supposed to be, single and unitary beings and if so, what is the principle of this unity. This can be called the problem of the unity of persons or, simply, the Unity question. The second issue involved here concerns the nature of the identity of persons through time - what makes (or makes us say, as we will maintain later on) a later person to be the same person as any earlier person. This problem has been described by some as the "Identity question" <sup>2</sup>; but I shall refer to this as the Reidentity question, so as to keep it distinct from another area of "identity" with which the present problem is not concerned.<sup>3</sup> The existence at all of these problems is due to the fact, noted above, that all observable features which we call the features of a single person are subject to change and so one is tempted to ask, what makes (us say) these features the features of one person rather than those of many different (or perhaps, succeeding) persons? Perhaps this is also backed by the argument, often advanced, that even the inner states of consciousness and experiences of a person are logically distinct and that some of them could occur to one without the others occurring at all. This has supplied the rationale to the question: what makes a set of experiences the experiences of a particular person rather than some-one else's? Because of this reason the unity problem has often been described as the problem of co-personality of experiences.<sup>4</sup> In order not to be presumptuous I prefer to call it the problem, simply, of co-personality; for the phenomenon of change and the fact of logical distinctness, as described above, can and do apply equally to the other (physical as well as psychological) features of persons, and to describe the problem of Unity as that of co-personality of experiences (alone) would, I think,

be tendentious in that it might lead us to be already working within the confines of a sort of Cartesianism. However, whether we take the problem as that of co-personality (of experiences and other features) or as that of reidentity (of persons), the problem is seen to be rooted in the supposed paradox, noted above, that everything about a person is subject to change and yet that the person is believed to be the same. How can a thing go on changing and yet remain the same? In other words, the question arose out of the view that the ideas of sameness and change are incompatible\*.(But for this the question of unity or co-personality of experiences would seem to be of questionable relevance to the problem of personal identity. As I see it, the only relevance it seems to have is due to the fact that it anticipates some unchanging principle, being in relation to which the different experiences are supposed to belong to the same person, and this 'unchanging something' was then easily supposed to be the bearer of personal identity. We shall shortly see, though, that this supposition was mistaken). The problem for those who fell into this way of thinking, then, was, what it is that remains unchanged about a person that makes them the single beings they are supposed to be and makes them the same throughout (what is called) their life-time? And this 'something unchanged', if any, was naturally to be other than the observable physical and psychological features\* which are known to be ever changing. The question, thus expressed in terms of a 'something unchanged', had the pretensions of addressing itself to the question of personal identity in both of its aspects of unity and reidentity. For if there is something unchanged which makes the different physical and psychological features belong to a single person then, it was naturally thought, it is the persistence of

---

\* Let us take "psychological features" in the wider sense to include the making of memory claims and displaying particular types of personality and character, which are observable phenomena. This is compatible with saying that remembering and having particular types of character and personality may be private and unobservable - a different matter.

that thing which would make a later person the same person as an earlier person. Such was thought to be the importance of this unchanged something that in the absence of a proof of its presence, not only were persons described as mere bundles of perceptions", the identity ascribed to them over time was dismissed as only a "fictitious one". But an enquiry into the nature of this "something unchanged" would make us see the dubiousness of such a supposition. The concept of a substantive self or soul has been often invented to meet the requirement of this "something unchanged". It was supposed, on the one hand, to constitute the essence of a person by virtue of which the different experiences( and other features) of a person are called or become his experiences and features, and, on the other hand, it was thought to be the bearer of his personal identity. However, the concept of a soul, as thus understood, though widely entertained, has hardly passed for an intelligible concept. Hume, among others, pleaded the privilege of the sceptic as regards its nature and existence. But whether or not the soul exists, and whatever emotional appeal and intimate feel it may have, it, being essentially private and unobservable, cannot make us say that a person is the same person as the earlier person he claims to be or as the person we think him to be. It cannot, i.e., be our justification for making identity judgments about persons. That it cannot be our justification for making such judgments about persons other than ourselves is fairly clear. For all that one can count on when one says that a person(other than one-self) is the same or is not the same as any earlier person is the conjunction or disjunction of certain physical and/or psychological features or continuities which are observable phenomena. And obviously none of these phenomena can be the soul since not only are they observable but are changeable which the soul is claimed not to be. It will not do to say that the conjunction or disjunction of those features are due to

the presence or absence of the soul which nevertheless remains unchanged while these features (its features ? ) go on changing; saying this would not be solving the problem at issue but only swapping it "for one of exactly the same form but less tractable"<sup>5</sup>

For our original problem was: what makes a person the same person as an earlier person when he changes radically in respect of observable features ? and now the problem would be replaced by : what makes a soul the same soul when it has changed so radically ? Only the problem has now become the problem of identity of something less tractable since persons are usually - unless defined as souls or as something non-physical - supposed to be observable. It is, thus, the difficulty of knowing the soul, if such is there, that makes it fall far short of being our justification for saying that a person is the same person as the person he claims, or is claimed, to be. Perhaps, this difficulty is not encountered in the case of the identity of oneself. It is often claimed by advocates of the soul-theory that the soul is known, in one's own case, by means of introspection, and that it is this knowledge that gives one the strong conviction that one has of one's own identity. This conviction was referred to by Reid as the conviction that "needs no philosophy to strengthen it and no philosophy can weaken it either"<sup>6</sup>. Well, this may be so, but what is not obvious is that this conviction is justified by the fact that the presence of the soul is known by means of introspection. At least it was not obvious to one person; his name was David Hume, who only "stumbled upon some perception or other" but never could "catch himself",

whenever he tried to introspect. What this proves is the fact that the knowledge of the soul by means of introspection is something not universal and that the fact that it is universal cannot be demonstrated. Now, even if it were true that the knowledge of the soul is the justification of one's own identity, its failure to be so in the case of other persons is enough reason for its unhelpfulness to the solution of the problem of personal identity. For it is in the latter case that there is a problem at all. There is, strictly speaking, no problem as to whether or not one is oneself the same person as the person one remembers being at earlier times. I do not need any justification for my identity through time. There is a problem about personal identity and a need to justify identity - judgments about persons when there is the possibility that both types of answers - an affirmative and a negative answer - are forthcoming to the identity question. Identity questions about other persons admit of both these answers - sometimes they are answered truly in the affirmative and sometimes truly in the negative, and even when one such question is answered truly in the one way the possibility of answering it in the other way is still open, and makes perfect sense. But in one's own case, by contrast, it does not make good sense - it would indeed be self-stultifying - to say that one is not the same person as the person "he" used to be five years back.\* And the answer to the identity question in one's own case is

---

\*We may ignore the fact that one can and does say of oneself that one is a changed person now, and the like. For, as will be generally agreed, ~~is~~ this way of talking refers not to the cessation of a period (followed by the appearance of another person), but to the change of personality - which certainly is a different matter.

usually always in the affirmative. Hence the problem at stake is not about one's own identity (I shall reserve the expression "self-identity" to refer to this identity hereafter). This is not to deny that one can, in certain peculiar circumstances, wonder if one is the person one thinks oneself to be i.e. one can wonder who one is. As a result of amnesia or total loss of memory I may well wonder whether I am P.K.M. But who I am will, in that case, be determined, not by me but by others so that the problem here is strictly not one of self-identity but has the logical status of the other identity issue. I would, in this case, be depending on the testimony of others who, in their turn, would be using the evidence of my bodily and/or psychological features in telling me who I am. And so, the "knowledge of self-identity" which I would thus be acquiring would be logically indistinguishable from that of other-identity; for others would be following the same method in telling me who I am as I would in saying who they are. Thus, if this (other-identity) is virtually all the problem is about and if the soul-theory of personal identity is no solution to this, then it follows that the concept of the soul is not particularly suited to the solution of the problem at stake. Nor will it do to say that it suits the case of what I have called self-identity alone; for that would be tantamount to saying that one means two different things by saying that a person  $\underline{P}_2$  at time  $\underline{t}_2$  is the same person as the earlier  $\underline{P}_1$  and  $\underline{t}_2$ , according to whether or not  $\underline{P}_2$  is oneself.\*

The soul theory of personal identity owes its existence and plausibility largely to the view, as we have seen above, that the ideas of sameness and change are incompatible, and to the consequent belief that what we are identifying, or supposing the presence of, when we say that a person

---

\* I owe this particular argument directly to Professor Swinburne.

is the same as some earlier persons is something that has remained unchanged though other observable things about the person have gone on changing. We have seen the difficulties of the concept of a soul in meeting the requirement of being the justification for making personal identity judgments. Our discussion pointed to the fact that the soul's failure to meet this requirement was due primarily to its being unobservable since, being so, it could not be shown to be unchangeable (or even to be changeable). But the failure of the soul-theory in the described respect need not oblige one to deny that there is such a thing as personal identity. The theory that persons remain the same through time is not incompatible with an outright 'no-soul' doctrine. One may totally reject the idea of a substantive soul or self and yet, pace Hume, believe in the identity of persons over time. Can it, then, be that the latter is because of the presence of some feature of persons, yet undiscovered and other than the soul, which remains the same when a person is said to be the same and is different when persons are said to be different? Such indeed would be a theory acceptable to those who do not believe in the soul and yet do want to maintain that something must remain unchanged in order to account for personal identity. But this secular version of the soul-theory fares no better than its original counterpart. This theory can only succeed by establishing the existence of an unchanged feature of persons which can be observed. But it will be a bad scientific hypothesis to expect such a thing to exist, for it is a general fact of nature that every phenomenal thing is subject to change.

Moreover, what seems to be a likely error of the theories of the above type is that, side by side taking the ideas of sameness and change to be incompatible (an error which will be exposed shortly afterwards) they have treated the questions 'what is personal identity?', 'what makes a person the same person as an earlier person?' as if they were straightforwardly

empirical questions. Being strongly motivated against something being changeable and remaining the same, the inventors of these theories have, perhaps, cherished the idea that there is something yet undiscovered which, if it could be discovered by means of laborious scientific investigation, would answer our question. But that the present issue is not straightforwardly empirical\* in this sense, and cannot be so decided is strongly suggested from the following consideration. If it were a straightforwardly empirical question in the described sense, then the appearance of unfavourable facts would have disinclined - if not logically compelled - us from applying the expression "same person" to the cases to which we do apply it. But we have seen that enough unfavourable facts are already there though they have done nothing to shake the conviction with which we use the expression and treat persons as continuous beings. It is not that we treat persons in this way because of any ignorance of facts; we do so treat them despite our full knowledge of the facts - which is that persons are "constantly changing and over a period of time may have changed completely" in respect of all their observable features.

The look the present question has of being straightforwardly empirical is due, perhaps, to the fact that the question is often asked in the form of a "what is ...?" question that is typical of the philosopher's way of asking questions. For, thus asked, the question does not always make it clear as to what is being asked. There are at least two senses in which a "what is ...?" question may be taken. It may be taken to be a question about the word or phrase ("personal identity" or "same person") or it may be

---

\*By our denial that the present question can be straightforwardly empirical should, however, be understood nothing more than the claim that this question is not to be decided by discovering some further facts about personal identity than what already are available to us. In particular, this must not be taken to mean that our question is a non-empirical or verbal one. Quite to the contrary, we do maintain in this thesis that the empirical fact that persons do display similarities in respect of physical and psychological features over time is highly relevant - indeed indispensable - to the solution of the issue.

taken to be a question about the thing the words stand for or signify. When asked as a thing question it may often be considered a boring enterprise to answer the question by pointing at the thing or phenomenon in question (though this is one of the ways of answering this question, e.g. when we point at a flash of lightning in the sky in answer to a child's query 'what is lightning?'); what is demanded is often a great deal about the nature, function and cause - in a word, explanation - of the thing or phenomenon (e.g. an explanation like, 'lightning is a form of electric discharge caused or generated by such-and-such factors under such-and-such circumstances, etc). In any case "what is ...?" questions, taken in the second sense are straightforwardly factual or empirical questions which can be (and are supposed to be) answered by referring to discovered (or discoverable) features the thing in question may have. That our question about personal identity is not an empirical question in this sense has just been shown. Is it then a question about the words "personal identity"? i.e. is it a question concerning the meaning of "same person". Here again, there is an essential ambiguity which must be considered in order that we may understand what is being asked.

There are various ways in which meaning-questions can be taken.<sup>7</sup> But for our present purpose it will be worthwhile to consider two senses. First there is the rather trivial sense in which the question "what is (or what is the meaning of) personal identity"? can be taken to be a question about the literal meaning of the phrase "same person". But clearly, the problem does not involve a request for an answer to this. That we know this is evident from the familiar fact that we apply the phrase to the cases of the right sort (e.g. I and the P.K.M. you saw last week) and that we do not apply it to the wrong sort of cases (e.g. P.K.M. and the Professor of Philosophy at the University of Keele). Secondly, there is the strict sense in which a meaning-question can be taken as a request for a definition

so that our question can be answered only by supplying a definition of the phrase "personal identity". In this sense, an answer to our question can be given by specifying a phenomenon, or a set of phenomena, in the presence of which - and only in its presence - the phrase "same person" can be applied. It will be argued in this chapter (and substantiated later on) that no strict and satisfactory definition of 'personal identity' can be given. By a "strict" definition I take it to be one that obeys the conditions just outlined, by a "satisfactory" definition I mean one that is not trivial. For example, if the word "same" is defined as "not different" (see for example, the Concise Oxford Dictionary) that is an extremely trivial<sup>8</sup> definition and fairly worthless as a definition. For it explains nothing since, if we follow this method of defining words, "different" would only be defined, and equally trivially, as "not the same". Similarly, if "personal identity" is defined as the "sameness of a rational being"<sup>9</sup>, this will be a trivially worthless definition in the same sense, and more particularly so when "rational being" is synonymous with, at least implied by, the definition of "person" (i.e. "thinking, intelligent being..") even by the same author. (And is not Reid's claim that "continuous uninterrupted existence is ..... necessarily implied in identity"<sup>10</sup> - if taken as a definition of identity - a trivial definition in the above sense?) Thus, if any definition has to do the trick it has to be a non-trivial definition of "personal identity" which I think is not possible. However, I will argue for this only after examining some attempts at defining this concept.

A very general way of expressing a definition of "personal identity" may be to express it in terms of the same body and/or of same memory and character etc. The immediate difficulty with such a definition is that we cannot, strictly, speak of same memory and character etc but only of (qualitatively) similar memory and character, which points to the fact that they are different and changeable over time. "Same body" is more in use

and is used to signify the numerical identity of a particular body that a person is supposed to have. But we have seen that what is called the same body is also subject to change and cannot, it might be argued, be called the same. It might thus appear that any attempt to define "personal identity" in terms of the sameness of body and/or memory and character etc would virtually result in defining the concept out of existence. For if the definiens is subject to change, how can the definiendum be the same? This is in fact the paradox, referred to earlier, that would not allow a definition of "same person" to succeed before it is itself disallowed. However, the difficulty - if it is a difficulty - is not so much a difficulty of defining "personal identity" as it is of finding out a same something in terms of which a definition could be formulated; and the latter difficulty is due to the fact that the ideas of sameness and change are incompatible. We have already seen that although change in the relevant sense is an undeniable fact, yet the problem of personal identity (and that of identity in general, as we shall see later in this chapter) is not one of finding out something unchanged to justify our practice of making identity judgments about persons. It seems therefore that the belief that "sameness" and "change" are incompatible ideas must be ill-founded.

What seems to have provided an inspiration to this view is a superficial lexical fact that one of the meanings of "same" (in some dictionaries) is "not changed"; and since dictionaries provide explicit definitions (i.e. definitions in terms of synonyms), this fact might have created the illusion of thinking that "same" and "not changed" are synonymous with each other. But although "not changed" may be one of the meanings\* of "same", it would be a bad dictionary indeed that treats the two expressions as exact synonyms. For in actual usage the idea of change is allowed for, and

---

\*More will be said about this in the next section of the present Chapter.

incorporated in, most substantive concepts. It is of course true that in some cases those two ideas yield a contrast as, for example, in the two conjunctions, "same note" and "succession of different notes" wherein the latter, involving change of notes, is incompatible with the former. But there are also hosts of other cases where obviously there is no such contrast; as, for example, "same tune" and "succession of different notes", wherein the former, far from being incompatible with the latter, is understood in terms of it<sup>11</sup>. Indeed "the same throbbing feeling" (a mental state) might consist in the succession of different mental states. Hence to insist, on grounds of apparent opposition between "same" and "different", that a mind cannot be the same mind at different times because it is a succession of different mental states or that a body cannot be the same body at different times because it is a succession of different physical states is to ignore the importance of the use of substantive concepts only in conjunction with which "same" and "different" have a use at all. This important fact was recognised by Locke who claimed that the concept of identity per se is incomplete the that what makes us say that a given entity is the same depends upon what sort of entity it is; and in recent writings the increasing importance of this fact is certified by the emphasis on the "same what?" question.<sup>12</sup> The dogma that sameness and change are incompatible is thus exposed and the fly is shown the way out of the fly-bottle; there is no longer any room for misgivings against ascribing unit and across-time identity to persons and things despite the changes that so obviously but innocuously infect them, for as we saw, certain changes are incorporated in many (perhaps most) substantive concepts. And our concept of a person is one such substantive of which it is perfectly harmless to say that it applies to a single being which continues to be the same at different times despite the described changes infecting all his observable features. Ascription of across-time identity to persons will not be incoherent - not certainly on

this score - though what changes are allowed for, and what not, in order for the sameness of a person to be retained is a further issue. It will be generally agreed that not all types of changes are allowed for by the concepts which do incorporate change into their meaning. For example, my bicycle may still be the same bicycle if some of its parts are replaced by new ones, but what if all its parts are replaced? Perhaps it will still continue to be the same bicycle if all its parts are replaced gradually over a period of time, but what if the parts are replaced all at a time? Similarly, a person may continue to be the same person with one or, perhaps, several parts of his body lost or replaced by "foreign" parts by means of transplants or plastic surgery. Very few, though certainly not all - would doubt that he won't be the same person if all (literally all) the physical parts of his body are replaced by plastic surgery for in that case we would not be having a person in the first place. But what if what is counted (because of its causal role in the exercise of memory and consciousness) as the most vital part, the brain, is replaced by another? What, in particular, if after the replacement of the new brain the person displays not only a total lack of "old" memories and character etc but also the exercise of totally new ones? Further, we accept a person to be the same person despite the alleged changes on the ground that the changes are "slow" and "gradual"<sup>13</sup>; but what if the changes are abrupt and total in some cases?

Thus, the problem of identity of a substantive turns out to be the problem of specifying what type(s) of changes are to be allowed for in order that across-time identity may be truly or justifiably ascribed to things coming under the substantive. And what changes (and what not) are to be allowed for depends on the nature of the concept and is largely a matter of specifying the criteria for the identity of things in question. It certainly is not the problem of how at all to ascribe sameness to

anything since everything supposedly is subject to change. This latter, it should by now be clear, is a spurious problem since the assumption on which it is based (i.e. that sameness and change are incompatible) is seen to be just mistaken. Consequently, the idea of a substantive soul - or, for that matter, anything unchanged - which was invented to answer this question can now be seen to be an unnecessary hypothesis, designed to solve a problem that does not exist. Now that the facts of the matter have been laid bare, there is for us no need of finding out some unchanging thing in terms of which to define 'personal identity'. All that matters is, in what sense the features of persons - changeable as they are - can be of help in giving an account (defining or otherwise) of personal identity. Although I want to maintain that the problem is one of specifying the criteria for the identity of persons, it is nevertheless useful and instructive to consider if any attempt at defining personal identity succeeds.

Any attempt at defining personal identity, in order to get started, must take into account one or several or all the features that characterise a person. This will not mean any - even incautious - step towards blurring the distinction between being a person and being the same person. But unless some, if not all, features of being a person are included in an account of personal identity we will not know what is being judged to be the same. For as we have seen "same" is incomplete and cannot be used in the first place, without a substantive concept conjoined to it. And since it is not obvious which, if any, of the characteristics of persons will do the trick, there seems to be no alternative but to examine the nature and role of all the characteristics (that seem to be essential to being a person) in defining or giving an account of personal identity. The characteristic features of a person fall into two broad categories: bodily or

physical and mental or psychological. Each of these sets of features is subject to change - though we have seen that this need not deter or detain us. It will do the trick if the changes are suitably qualified so as to enable us reasonably to make personal identity judgments on the basis of these characteristics. The qualifications that are usually made is that the changes in the mental and the physical features of persons are slow and gradual and uninterrupted. This is otherwise expressed in the notions of bodily continuity and the continuity of memory and character. These continuities, jointly or severally, can and do serve as the justification for making identity - judgments about persons. But though these continuities are our justification for ascribing identity to persons, it will be presently seen that any attempt to define "personal identity" in terms of these continuities can hardly succeed.

There have, in fact, been certain theories\* which have apparently taken the problem as a problem of definition and have tried to define personal identity in terms of the one or the other or both of these continuities. The most acceptable and workable interpretation of these continuities has been in the following manner: bodily continuity has been interpreted to include spatio-temporal continuity such that a body  $B_1$  is the same body as the body  $B_2$  if they are connected by a spatio-temporal path at each point on which there is a body which is somewhat similar - in appearance and constitution - to its immediately preceeding and succeeding bodies which are on the same spatio-temporal line. Continuity of memory and character has also been spelt out in a similar manner. Since memory and character cannot be conveniently talked of without reference to the persons whose memory and character they are, it is best to state the nature of this

---

\*Professor Swinburne calls them the "empiricist theories" in a paper "Personal Identity" Proceedings of the Aristotelian Society (PAS hereafter) 1974. The present chapter owes a great deal to this paper.

continuity via what according to the memory theorists constitutes the sameness of a person. Thus,  $P_2$  at  $t_2$  is the same person as  $P_1$  at  $t_1$  (assuming there to be a considerable time-gap between  $t_1$  and  $t_2$ ), "if there is a series of persons  $P_n$  at times  $t_n$  intermediate between  $t_1$  and  $t_2$  such that the memories of each person include almost all those of any person slightly earlier in the series and each person is very similar in character to any member in the series existing at a temporally proximate moment".<sup>14</sup> With these sort of interpretation of the two continuities, the empiricist theories have felt comfortable to define personal identity in terms of the one or the other or both of these continuities. Two rival theories of personal identity have emerged as a result of emphasising one of these continuities to the exclusion of the other. The relative merits and demerits of these theories we will have occasion to judge in later chapters. But as attempts to define personal identity in terms of either or both of these continuities the weakness of such theories will be fairly ~~manifest~~ manifest.

I have said earlier that something can be the definition of something else if in the presence of the former, and only in its presence, the latter can have application. Neither bodily continuity alone nor memory-and-character continuity alone is competent to supply a definition of personal identity in this sense: the belief in a person's survival of his (bodily) death and continuing in a disembodied state, though less common, is not logically impossible, and equally possible is the idea of reincarnation.\* The former points to the possibility that there can be personal identity without there being bodily continuity and the latter to the possibility that there may be no personal identity despite there being a continuous body which might possibly be "inhabited" by a different person as soon as or soon after "deserted" by the original person. On the

---

\*I will not, however argue for these forms of personal survival and the later sections of the present thesis will bring out some serious limitations of these ideas.

psychological side, there are the fairly frequent and normal phenomena that a person should lose his memory and/or radically change his character and personality and yet remain the same person as before; and it is also possible that two or more contemporaneous persons may make the memory claims\* and display the personality and character that fit the life of a third, non-contemporaneous person and yet, for well-known reasons, this cannot be a case of these (contemporaneous) persons being the same person as they claim to be.<sup>15</sup> These actual and possible cases show that neither bodily continuity nor memory-and-character continuity can be a necessary condition of personal identity and also that neither is a sufficient condition. It needs scores of philosophical arguments to show that the possible cases referred to above are not cases of personal identity (or non-identity as the case may be), and so the foregoing should be enough to damage any claim to analyse or define personal identity in terms of (one/the other/both) these continuities.

Now because of what has been said above, an empiricist theory which may attempt to analyse personal identity in terms of both bodily and memory-and-character continuity will do no better either. In order for such a theory to succeed (i) each of the continuities has to be a defining characteristic and hence a necessary condition of personal identity and (ii) both together (and not one alone) would have to be sufficient condition. But each of these continuities has been seen not to be necessary, so condition (i) is violated. And as regards (ii), it has been shown that neither bodily continuity nor memory-and-character continuity is a sufficient condition of personal identity (by imagining respectively

---

\*It will not do to object here that 'making the memory claims of' is not the same thing as 'having the memory of'. For it is in terms of the former that a definition of personal identity can be formulated at all: to define it in terms of the latter would be inevitably circular.

the conceptual possibilities of a body being "deserted" by its original "occupant" and soon after being incarnated by another person and that a later person may start off by making memory claims (and displaying character and personality) similar to that of an earlier person who, on independent historical grounds, could be seen to have been a different person). Now, in order to see that both these continuities together cannot be the sufficient condition either, we have only to imagine a combination of these two possibilities, i.e. a person, A, may die and his death may be (it is possible to imagine) instantly followed by another person, B, having exactly similar psychological features as A's taking over the latter's body. And what this possibility will show is that even in the presence of both bodily continuity and memory continuity there could be no personal identity. Thus will condition (ii) be violated. It will follow that a definition of personal identity even in terms of both the described continuities together will not succeed. And from this, together with what has been said in the last paragraph, it will follow that no empiricist theory of personal identity can succeed. For a theory of that sort will succeed only by defining personal identity in terms of one or the other or both of the described continuities and this is an impossible task.

There has been another serious recent attack<sup>16</sup> on the efficiency of the empiricist theories on the ground that "because there are imaginable circumstances where any criterion (memory or bodily continuity, PKM) gives no clear result, there are imaginable circumstances where any empiricist theory of personal identity would give no answer as to whether  $P_2$  is the same person as an earlier  $P_1$ " and that "the answer that they are the same is as near to the truth as the answer that they are different"; and certainly, Swinburne rightly implies, we would not want to say a thing like that about personal identity. The "imaginable circumstances" are the numerous puzzle cases that abound in the literature of personal identity.

And this criticism further substantiates my contention that the empiricist theories cannot solve the problem of personal identity. (Moreover, if a theory claims to be a theory of definition, how can it succeed without the definition it proposed applying to all cases - real or imaginary? How else it is a definition?) My treatment of the empiricist theories makes it clear that the failure of these theories is due not to the fact that they count on bodily continuity and continuity of memory and character for the solution of the present problem - as we shall see, these continuities are perfectly alright as our evidence or criteria, and as a matter of fact these seem to be the only things we can fall back upon - but to the fact that they wanted to define personal identity in their terms, which we have seen to be impossible. And I cannot see how else it can be defined, since these continuities are the only notions in which our understanding of the concept of personal identity seems to consist. As Swinburne rightly points out,<sup>17</sup> we get to understand the meaning of the phrase "same person" by being shown clear cases of persons who are, and of persons who are not, the same; and in all such cases we are provided with the presence or absence of one or the other or both of these continuities or similarities. There seems to be nothing else which would acquaint us with the notion of personal identity; if there were any it would have been shown to us, - otherwise the meaning of the phrase "same person" could not be taught to us and no one could have learnt its meaning. But if there is nothing else but the continuities in question and since the latter are incapable of defining personal identity, then it seems to follow that the notion of "personal identity" cannot be defined. This is the element of truth in the claim, notably made by Butler and Reid, that personal identity is something ultimate and unanalyzable.<sup>18</sup>

As we have seen, a similar conclusion is reached by Swinburne. But the conclusion needs to be rather carefully interpreted. As Professor Mackie has rightly implied,<sup>19</sup> Swinburne's conclusion that personal

identity is unanalyzable goes hand in hand with, and indeed issues from, the idea that our ordinary concept of a person is "something whose unity is in itself unequivocal" (ld p 193). And he, like me, agrees with Swinburne (and quotes him with expressed approval) that the observable continuities delineated above are the only means of our understanding and coming to know the meaning of personal identity, but that these continuities are not what we mean by "personal identity". Yet Mackie complains that Swinburne's conclusions that personal identity is something ultimate and unanalyzable does not follow. By examining the status of those continuities, severally and jointly, and by showing that there is nothing else in terms of which personal identity could be defined, I have argued that the conclusion is inevitable. What, then, is Mackie's problem? If I have understood him rightly, his problem is absolutism: he believes that the Butler-Swinburne type of theory is committed to the belief in a spiritual substance whose persistence must account for the identity of a person through time.\* This sort of commitment is not entirely unusual; as I shall argue in the next chapter, Butler, at least, expressly espoused such an assumption. But I am not sure if such a commitment is logically binding. Swinburne, in a way, implicitly contends that the relevant continuities are all there is to our understanding of personal identity ("of what is at stake" op cit, p241), and also that the latter is "observable only by observing these" (op cit, p 240). But saying this does not commit oneself to saying that we observe two things: these continu<sup>ies</sup> and personal identity which is the continuity of some spiritual or non-physical or absolute substance. To take the help of an analogy, we observe the team spirit of Stoke City Football Club by observing the "potters" playing; but that is

---

\*See esp Thid p.194 "Have we not come back", Mackie asks, "precisely to the notion which, dressed up in the philosophical terminology of spiritual substances, Locke so rightly rejected and set aside?"

not to say that we observe the former besides observing the latter (though we may still insist that "team spirit" means something more than the players playing in the way they do). In view of this Mackie's problem of how the ultimate identity of persons can be observed only by observing the relevant continuities need not be worrying at all, nor is it anything "obscure" to say this, nor is Swinburne bound to say that the identity of persons is the identity of something ghostly or spiritual\*. It can be said with perfect propriety that a table is observable only by observing all its observable features, and yet that "table" does not mean only the observable feature of a table; but saying this does not entail that table is some unknowable material substratum which is what we mean by the word "table". I conclude, therefore, that the belief in a spiritual substance does not inevitably follow from the theory, which Swinburne maintains and I subscribe to, that personal identity is indefinable. The belief, though, in a spiritual substance, instead of following from the contention that personal identity is unanalyzable, will lead to the latter - which, I believe, is Butler's way of reasoning, at least in part. On the contrary, I have maintained (see pp7-8 above) that this belief in a spiritual substance issues from the insistence that there must be something unchanged to account for our ascriptions of identity to persons. We have seen this to be a mistaken belief and that it is unnecessary to the belief in, and ascription of, identity to persons. The fact rather is that we talk of personal identity and make personal identity judgments on the basis of bodily continuity and/or the continuity of memory and character etc, but that we also can, and do, meaningfully ascribe identity to persons even in the absence of one or the other of these continuities. This much is

---

\*Perhaps the use of the expression "ultimate" is slightly misleading here; but this expression may very well be understood as a synonym for its accompanist "unanalyzable" - as indeed Swinburne tells me that he meant to use it in this way.

the spirit of the claim that personal identity means something more than those continuities and so cannot be analysed in their terms. And since these continuities are all that is there to our understanding the meaning of personal identity, it follows that the latter is indefinable.

The offshot of our discussion has been that the question "what is personal identity?" cannot be taken either as a question about the literal meaning of the phrase "same person" or as a question about the definition of personal identity. For while the former would be unnecessary (because every user of our language knows how to use the phrase in its normal context and hence is in no doubt about its literal meaning), the latter is seen to be unfruitful (because of the impossibility of defining the notion). This difficulty can be expressed by saying that though we know what it means for a person  $P_2$  to be the same person as an earlier  $P_1$ , we cannot say (cannot state in the form of a definition) what personal identity is. In other words, although we know, and can say, when two non-contemporaneous persons are, or are not, the same, we cannot say what exactly makes them the same person. In the face of this difficulty, the only important way of investigating the nature (which by no means is giving a definition) of personal identity would be to consider how we know that two non-contemporaneous persons are (or are not) the same. We know that a person is the same person as an earlier person on the basis of evidences, and bodily continuity and/ or the continuity of memory and character etc are our evidences. Under all normal circumstances the conjunction and disjunction of these continuities help us (are our justification) in deciding which cases are, and which not, cases of personal identity. In different types of situations judgments of personal identity can be made, and in fact are made, without much difficulty by carefully balancing the evidences against

each other. But if the situations are sufficiently different, or are imagined to be so different, the problem of how to make identity-judgments becomes more serious and presents more serious difficulties; the question then becomes not what are our evidences, but what are the "criterial evidences" or, simply the criteria for making judgments of personal identity. Different interpretations have been given of the notion of "criterion" and I will discuss them and suggest mine in Ch 3 Sec. 1. For our present purpose a criterion can be roughly described as a state of affairs the existence of which is always or necessarily evidence for the truth of the judgment of which it is a criterion. In sufficiently different situations, therefore, such as the so-called puzzle cases, the questions of personal identity are particularly about these evidences which are criterial for the truth of personal identity. It is these situations which raise the problem more conspicuously than ordinary situations (the problem hardly ever occurs to us in ordinary situations); I believe that most of these situations help raising the problem in a clearer perspective, since it is precisely with this motive that the puzzle situations are brought into existence in the literature of personal identity. And it is the existence, or at any rate the possibility of such situations which, I think, brings the truth to light - which is that our problem is a problem of specifying the criteria of personal identity - and not a problem of defining it, since these possible situations most palpably resist any attempt at defining in terms of any prospective definiens available to us. Thus Shoemaker: "The problem of self-identity", (by which he means personal identity), "is often characterised as the problem of specifying the criteria of personal identity."<sup>20</sup> The word "often" in this quotation is misleading since it may suggest that on some occasions the problem may be not about criteria, but possibly about the meaning or definition of "personal identity". But

taking the problem in this sense, we saw, is either unnecessary or unfruitful; and if one is not clear about this, one would be confused even before beginning to find any solution. My contention therefore, is that our approach to the problem would be on the right lines if it is taken as a problem of specifying what criterion or criteria should be used to make and judge statements of personal identity. Shoemaker notes that in recent discussions of the literature the questions "How is the identity of  $\phi$ 's known?" and "In what does the identity of  $\phi$ 's consist?" are often reduced to the single question "what are the criteria of  $\phi$ -identity?".<sup>21</sup> This is true particularly of our contention about the nature of the problem of personal identity, though this remark, being a remark about the general problem of identity, leads us on directly to consider the distinctive feature, if any, of the problem of personal identity, as against the identity of other material objects. This will be our concern in the following section.

## Section (ii) - A Special Problem

The problem of personal identity, as it has been stated in sec (i), and as it has been generally considered therein, is a problem that can as well be raised about the identity of material things. This may have the effect of lending support to a not very uncommon view that persons are not essentially different from mere material things and that there is no essential difference between personal identity and the identity of those material things.\* There seem to have been two possible reasons for the last view to be plausible. One reason, which makes only a negative point and which concerns the nature of identity, seems to have been that invariance is often thought to be the standard of identity or of sameness in all cases. We have seen that some philosophers - most notably Hume - have insisted that in order for anything to be called the same it has to remain unchanged from one time to another. For these philosophers, "being the same" and "being unchanged" are synonymous. Thus, after having reasoned that everything that is supposed to last beyond a moment is a "bundle" or "collection" of rapidly changing perceptions and hence is not entitled to be called the same at different times, Hume says that "the same method of reasoning should be continued" to show that the identity ascribed to persons (which he calls 'the mind of man') "is of a like kind with that which we ascribe to vegetables and animal bodies". In other words, identity is to be ascribed - or rather denied -

---

\*The expression "material things", as it will be used throughout, will be understood to mean not only tables and stones but also animals and organisms - to include, in fact, all observable things except persons. This will, thus, be a very broad use of the expression "material things", and as our treatment of "person" in this section would indicate, the latter will also come under that description. So that when I consider persons vis-a-vis material things, the latter strictly should be understood to mean other material things. I do not, at any rate, intend to suggest that persons are non-material or non-physical things, and nothing that will be said about the distinction of persons from "material things" would commit us to that idea. My purpose simply will be to try and show that the nature and identity conditions of persons differ in important respects not only from those of mere material things like tables and stones but even from those of animals. Furthermore, although the arguments in following will be more often concerned with showing the relevant contrast between persons and mere material things, a subtle distinction will also be indicated between persons and animals.

to both persons and other material things in the same sense. For the ultimate outcome of Hume's reasoning was that the identity which we ascribe to persons as well as to material things is "only a fictitious one".<sup>1</sup> Of course Hume's thesis stands or falls along with his basic assumption that things that change cannot be really the same, and we have already seen that this thesis is mistaken. Penelhum has argued that Hume's error lay in thinking that invariance is the standard of identity in all cases whereas, in fact, it is the standard only in a few cases.<sup>2</sup> This criticism draws force from the way the expression "the same" is used. In most cases it is used despite the object undergoing changes (e.g. 'the same bicycle', 'the same church', and the like). We have seen that in such cases changeability is incorporated in the concept under which the object in question is grouped. In the few cases in which "the same" is used in the sense of "being unchanged" are the cases where unchangeability is a part of the meaning of the concept. For example, we say of a musical note that it is the same note only when it has not changed; if there is any change, we have a different note but no longer the same note. But if this is so, that is because the concept is defined, partly at least, as unchangeable. But then the statement "It is the same x (where 'x' is defined in terms of unchangeability)" is not an identity-statement in the sense in which we are considering such statements, and the 'sameness' (and 'difference') of x is irrelevant to the problem of identity. For we have said that the identity with which we are concerned is concerned with the reidentification and persistence of things through time and change. In this sense, a statement like 'I saw the same flash of lightning as the one you saw' does not even imply identity, though it is an identity-statement in another sense (see note 3 to the previous section). It follows therefore that "being the same" and "being unchanged" are not synonymous which they have to be in order for Hume's thesis to stand, and

that where they are synonymous the "sameness" is irrelevant to the problem of reidentity. Consequently, Hume's claim that 'identity' has the same sense (i.e. of being only a fictitious one) when applied - if applied at all - to both persons and other material things loses ground. And if the supposed synonym was Hume's ground for claiming that identity as applied to both these cases "is of a like kind", then he was clearly mistaken.

However, our criticism of Hume has not shown what, if any, is the difference between personal identity and the identity of material things though it is common to feel that there is a difference. But something has emerged from this criticism to the effect that invariance is a part of the meaning of certain concepts and not of others and, hence, that in what sense "the same" is used would depend on the thing to which it is applied. This principle, though extremely crucial to the issue at stake, may not seem to be particularly helpful because of the way it has issued from the foregoing. For it was shown that where invariance is a part of the meaning of the concept sameness and difference of these sorts of things are irrelevant to the problem of identity so that it might (which though, I hope, is very unlikely) be thought that since there are two sorts of things, one that does not allow changeability into its concept and the other that does, and since it is only in the latter case that 'sameness' is relevant to the problem of identity, then "sameness" has the same sense when applied to this class of things. But we have already said in the last that although change is allowed for by most of our concepts yet not all types of change are allowed indiscriminately by all these concepts; and what changes will be allowed for a thing to be the same thing depends upon the thing in question. This, in fact, is the spirit of the principle that emerged - or, rather re-appeared - this gives us an important clue to understanding the supposed distinction by concentrating on the nature of the things to which

we apply the words "same" and "different". But before that, let us consider the second reason which seems to have led to the idea that there is no distinction between personal identity and the identity of material things.

This reason, which concerns the nature of persons and material things rather than that of identity, is that there is a sense in which persons are material things. Certain things are said about material things which are also said about persons. For example, we say of a person that he is such-and-such feet high, weighs such-and-such pounds, is located in physical space and the like. One can even say though it may be slightly unnatural and strenuous way of speaking, that a person is soft or hard (not meaning soft-hearted or hard-minded). There is, further, the apparent unknowability - of "other minds" which has a tendency to make us treat behaviouristically somewhat as we treat the ordinary material things. We also say of a person such things as: he was removed to the hospital, that I pushed him away etc, which are analogous to: the table was removed to the next room, I pushed the chair away and the like. Such analogous ways of talking about both persons and material things make it appear that there is no essential difference between the two and, accordingly, that the way we make and judge statements of personal identity cannot be any different from the way we make and judge statements of the identity of material things. Thus Reid for example said "Our judgments of the identity of the objects of sense seem to be formed much upon the same grounds as our judgments of other persons beside ourselves."<sup>3</sup> (my emphasis). Reid's remark is revealing of two different facts. On the one hand, it points to the fact that there seems to be no difference between persons and material things (which, in turn, was thought to be responsible for there

being no difference between personal identity and the identity of other things), because of similar ways of talking about both persons and material things and because of apparent lack of knowledge of other persons except in physical and behavioural terms. On the other hand, it suggests implicitly that there must be a difference - we feel that there is - in our own case, at least. Keeping this latter point for later consideration, it is not difficult to see why it is that we say similar things in both cases and how it is that saying similar things in both cases has done nothing to make (or make us think) persons the same type of things as material things. Obviously, we do say such things about persons as we say about the material things because persons have bodies which are material things after all. But if from this one is inclined to reason that persons are merely material things then one would be doing the odd job of deducing " $\phi$ 's are  $\psi$ 's" from " $\phi$ 's have  $\psi$ 's". Moreover, equally obviously, certain other things are also said about persons which are not said - which it makes no sense to say - about material things, and animals which we have agreed to discuss under the heading of other material things. For example, we say only of persons that they think, know arithmetic and solve philosophical problems. This is due to the fact that persons have minds; and in whichever terms the latter may be interpreted, this is what makes the psychological attributes (or what Strawson has called the P-predicates) ascribed to persons and makes persons distinguished from the other material things - in the broad sense which we have given to the latter expression. If a lion or a stone could do what only a person can do, we would try to attribute this somehow to a mind somewhere or, if we cannot possibly do that, leave it as something mysterious - which is certainly not to explain it.<sup>4</sup> The essentialness of having a mind to persons explains why it is that "what is a person?" and "what is a mind?", "In what does personal identity consist?" and "In what does the identity

of the mind of a man consist?" are often equated. As I have said above it is odd to deduce " $\phi$ 's are  $\psi$ 's" from " $\phi$ 's have  $\psi$ 's", and this is as true of the claim that persons are bodies as of the rival claim that persons are minds. Yet the distinctively essentialness of having minds to persons <sup>has</sup> ~~is~~ apparently camouflaged this oddity and has been rather over-emphasised to make Cartesianism or near-Cartesianism more plausible than in fact it is. At least one obvious error of the last mentioned theories seems to be that from "x is essential to y" they have, in effect, inferred "only x is essential to y".<sup>5</sup> But the latter, clearly, does not follow from the former; for may it not be that x is one of the essential properties of y? Being equiangular is essential for something to be an equilateral triangle, but it does not follow from this that only being equiangular will make an equilateral triangle. (Can a rectangle or a square be an equilateral triangle?) This is to forget - if only one can - that the latter is also essentially a triangle. There seems no prima facie reason why similar reasoning cannot be accorded to the case of mind's being essential to a person. For while this is no doubt true, it is not incompatible with something else - perhaps 'having a body' is that - being also essential to a person. Now, if the philosophical tendency to identify persons with their minds were justified it would not only show the essentialness of having a mind to persons, it would also show the body to be inessential. As a matter of fact this tendency has such natural and familiar appeal that though disembodied persons are often thought to be possible, mindless persons are hardly conceived at all. However, admitting that minds are essential (in being the distinguishing feature) to being persons does not commit us to such a radical position. The fact rather is that persons and material things

are both bodied,\* but whereas material things have bodies only (let us ignore the awkwardness of saying this - though "having a body" can strictly be said to have a different sense<sup>6</sup> when applied to material things from the sense it has when applied to persons), persons are distinguished by the fact that they have minds. As Strawson has so succinctly expressed this "among the things we ascribe to ourselves are things of a kind we also ascribe to material bodies to which we should not dream of ascribing others of the things that we ascribe to ourselves"<sup>7</sup> (emphasis added).

Now, an important consequence of the fact that persons have minds, which shows the distinction between persons and material things to be very fundamental, is the fact that persons are self-knowers. Not only are we persons, we know that we are. What it is to know that we are persons and what relevance it has on the nature of personal identity? Consideration of this question will lead us to the understanding of the nature and peculiarity of the problem of personal identity.

To know that I am a person is to know that I think, reason, intend, remember etc to mention only a few. Obviously, I am not though, not

---

\*For the present purpose, it will do, I hope, to bear in mind that it is a general fact of nature that persons are bodied beings and that we get to know persons (and the meaning of the word "person") as bodied beings. We say of such a thing (who, besides being whatever else, is a bodied being), and what is like one, that he is a person; and we are taught the meaning of "person" by being shown one or several of these beings. But for this, I cannot see how one can say that something (what things?) is a person and how the meaning of the word could be taught to anyone. Any otherwise account of persons will not be our concept of a person and will belong, at best, to some other area of discourse. This may not prove the impossibility (logical) of there being disembodied persons; but it certainly suggests that the latter concept of a person would be an extended (and so secondary) version of our normal concept. I shall return to this question in Ch 4 sec (iii) to show that this extension of the normal concept of a person is not justified.

reasoning and the like; I am something (or someone) which (who) thinks, reasons, intends and remembers. But I do not do these things only once in my life; I do these many times. None of these is continuous and uninterrupted throughout the whole course of our life. I have different thoughts, intentions, memories etc. at different times; even if on a number of occasions I have qualitatively the same (strictly similar) thoughts and intentions and memories, they <sup>are</sup> likely to be, interrupted by other thoughts, intentions or memories, in between these occasions. (It would be a very strange world if people there were having the same (numerical) thought or memory uninterruptedly for all times). Yet these different and many activities I call mine. It is not that I only call them mine, I know them to be mine. What makes me call them mine or what makes them mine is my continued existence or, if you like, the fact that these activities are performed by me who continues throughout the operations of all of them at different times. Moreover, as Reid has pointed out,<sup>8</sup> I cannot reason without it being true that the antecedents have been seen or done by me (by the one who reasons), nor can I remember anything without the conviction that I, the rememberer, existed when the thing remembered occurred. Thus, the knowledge that I am a person carries with it the belief about and the capacity to know my personal identity. (If this is true of me because I am a person and, hence, a self-knower, this is true of others as well, since others are persons and so self-knowers.) And this, in a very important sense, makes personal identity constitutive of the concept of a person. Locke for example, defined "person" as "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places"<sup>9</sup> (my emphasis), thus making the knowledge of one's personal identity - (self-identity) a part of the definition of a "person". As

Locke's definition of "person" is quite a plausible one, it is also plausible to say, as the definition suggests, that there is an intimate connection between the concept of a person and the idea of one's personal identity; that if there is a person he is capable of knowing of his identity through time. This cannot be true of other material things.

A stone or a dog (certain reservation about the latter follows shortly) does not know - makes <sup>it</sup> no sense to say that it knows - anything about itself or its identity across time. If it is argued that dogs and perhaps some other animals behave in ways which suggest that they know certain things, my first response is that that "knowledge" is too vague and dubious to be worth the name. Wittgenstein said something to the effect that though a dog can be said to know, e.g., that his master is at the door, no behaviour of it could possibly show or even suggest that he knows that his master will come the day after tomorrow.<sup>10</sup> Instinctive behaviours can hardly bear any title to knowledge. How far it will be anything unnatural to say that instinctive behaviours are more or less like natural signs? (Compare our saying that it is going to rain at the sight of lowering black clouds and our saying that the master is at the door by seeing the dog behave in a characteristic way.)

It might, however, appear that what the foregoing and what in particular the remark of Wittgenstein proves is that although a dog can be said to know certain things, it cannot be said to know certain other, more complicated, things. But if, for this reason, one wants to say that the dog does not know anything even where he can be said to know something then one would be throwing the baby away along with the bath water - as James had complained against Hume.<sup>11</sup> For it is admittedly true that some characteristic animal behaviour do strongly suggest that they know certain things at least, although  $\pm$  no behaviours of theirs are particularly suited to signify the possession of higher order knowledge by them.

I am prepared to concede this though I am in no fear that this concession will affect my position which is that persons are distinguished from animals and inanimate things('material things' - as I have called them) by virtue of the fact that they are selfknowers. And it can hardly be denied that self-knowledge is knowledge of a very high order<sup>12</sup>, which no animal behaviour is suited to testify. Lack of evidence may be no proof of non-existence, but the availability of evidence is certainly a privilege where two rival classes of things are concerned. The evidence of self-knowledge could be available in the case of persons(self-consciousness in one's own case and self-conscious behaviour in the case of other persons), and not in the case of animals - not to mention the inanimate things.

Let us now see how the results of the foregoing can be utilised to bring out the special nature of personal identity as distinct from the identity of material things. And since persons are distinguished by the fact of self-knowledge, the best way to approach the issue would be by considering the way personal identity is known and the way the identity of material things is known. As we consider this, it will appear once again that it is the nature and possibility of self-knowledge that makes all the difference.

We have seen that persons are self-knowers and that this means that they know of their own identity through time. But equally, they also know the identity of the many material things around them. What, then, is it that makes the former knowledge so peculiar and the problem of personal identity - not merely self-identity - so special ? To come to an answer to this question we will start with how we come to know of a material thing that it is the same thing as the one we saw earlier. We always count on certain similarities which we normally take to indicate its continuity in space and time. This means that spatio-temporal continuity

is taken to be the criterion of identity in this case. This criterion is always taken to be a necessary (and perhaps sufficient) condition of the identity of material things.<sup>13</sup> Once this condition is satisfied no further evidence is normally required. And with certain reservations (to be clear in course of what follows), this seems to be also true of the identity of persons other than ourselves. Under all normal circumstances we count on certain similarities namely, these of physical appearance and of memory and character (sometimes the former alone, sometimes both) which are normally taken to be there because of the spatio-temporal continuity of a human body. Doubts may, however, be raised as to whether and how memory-and-character similarity can be taken to be due to the spatio-temporal continuity of the body. To this it may here suffice to reply that memory is taken to be a criterion largely because of the fact that there is such a thing as real memory i.e. if all the memory claims that persons made were deceptive and if none of them ever pointed to the identity of the rememberer with the witness of the past event that is remembered then memory would never be a criterion of personal identity. It is this general fact that some (in fact most) memories are real and hence point to the identity of the rememberer with the witness of the remembered event that creates in us the natural <sup>presumption</sup> ~~belief~~ that any memory claim may be real memory and hence the person making it may be identical with the witness of the putatively remembered thing. But a real memory there could be only if the person remembering the thing was physically present when the event occurred. Similarly, if character continuity is ever taken to be a criterion of personal identity it is because persons do behave in similar manner and do display similar traits of personality over a period of time; but this means that persons are physically observed over a period of time (Shoemaker deals with this point more elaborately in his Self-Knowledge And Self-Identity - Ch 5 secs 9 and

10). If, on the contrary, it could be otherwise established that the body of the person was not present when the remembered event occurred then, even if the remembered event did in fact occur and the person's 'memory report' fits the latter fact with excellent accuracy, the identity of the person will be extremely suspect.

It may be objected that the foregoing does not show that these criteria (similarity of memory etc and of physical properties) cannot be used independently of spatio-temporal continuity. I am not sure if this objection could be true against the use of the physical criteria. The possible explanation may be found in the cases - if such cases are admissible - of disembodied persons. In such cases the spatio-temporal continuity (of the body) being ex hypothesi inapplicable the only usable criterion of identity would be the continuity (or similarity) of memory and character. Admitting the possibility and the plausibility of such cases, I have said already that these cases would be an extension of the normal cases and thus are parasitic upon them.\* But in the normal cases, we have seen, that the similarity of memory and character serve as a criterion of identity in so far as it is supposed to be the evidence of spatio-temporal continuity as is relevant in the case of persons. And this fact that these criteria do serve as our criteria of identity because they normally indicate spatio-temporal continuity <sup>might have</sup> ~~has~~ inclined many to think that personal identity and the identity of material objects is known in the same way.

Thus, we saw, some philosophers have said, that our knowledge of the identity of material things and that of the identity of other persons

---

\*But see Ch4 secs iii and iv for a fuller argument for this.

beside ourselves are based much upon the same grounds. We emphasised the qualification "beside ourselves" and promised to discuss its significance later. In the light of the further discussions thereafter, I think it will be proper and helpful to discuss this point. As we said, this phrase was apparently used by Reid to indicate that there must be a difference between our knowledge of personal identity and that of the identity of material things. But since there seems to be nothing in the way the identity of other persons is known to show any difference, the difference was thought to lie in the way we know our own identity. This has an important element of truth - as important as to stress the difference - and we will soon try to make it explicit. But let us cast a quick glance at the suggestion that there is nothing in our knowledge of other persons to show the difference, for this seems to be not only misleading but also an act of oversimplification. It is noteworthy that Reid probably did see that there is a difference here, although he was not able to say what it is. This can be read into his remark that our knowledge of the identity of material things and that of other persons are based much upon the same grounds, - he did not say (nor mean, I take it) exactly upon the same grounds; and I will try to show that it was not merely an accident or a slip that he did not say that. For at least one way in which the difference can be brought out is by pointing to the fact that although our knowledge of the identity of material things and that of the identity of other persons are apparently based on the same grounds, yet these grounds are normally decisive enough in the one case while they are hardly ever decisive in the other. This is explained by the familiar fact that while spatio-temporal continuity is largely accepted as the necessary <sup>condition</sup> ~~conclusion~~ of the identity of material things, its like status in the case of (other) persons identity is highly controversial in the literature of personal identity. And although this criterion is often enough in the one case,

it is often thought to be not enough in the other.<sup>14</sup> That is to say, if a certain material thing is observed to be continuously present from one time to another, it leaves us in no conceivable doubt that it is one and the same thing, and there is nothing else which could make this identity suspect. On the contrary, though the observed spatio-temporal continuity of a living human body makes us fairly sure that the person in question is the same throughout that period, yet there are certain other conditions (e.g. the absence of a possible incarnation of another person in that body) which, if not satisfied or taken for granted, would make the supposed identity open to doubt and may supply a reason for thinking that it might be a case of non-identity as in the putative case of "bodily-transfer". This, I should think, is a conceptual difference, a difference that issues from what 'person' means and what 'an (ordinary) material object' means. "The spatio-temporally continuous table is the same table" is an analytic statement, whereas "the spatio-temporally continuous body is the body of the same person" is not; and hence the spatio-temporal continuity of the body would not, unlike the other case, be a conceptual guarantee of personal identity. Thus it is that the very same ground is adequate, at least generally so - in the one case and yet is not so adequate in the other. But if this is so then it seems to follow not only that our knowledge of identity in each case, if based on spatio-temporal continuity, is different (well-grounded in the one case, but less so in the other), but also that there must be a difference between the nature of identity in the two cases. But since it is not easy to see what the difference is from the case of other persons, it seems natural to suppose that the real nature of personal identity is to be sought, not in this case but in one's own case. It is an admitted fact that our knowledge of the identity of material things (and of other persons) is always grounded on some evidence or criteria, and that we have to use some criterion or criteria to know that

a later x is the same x as an earlier x (where 'x' is an ordinary material object or a person other than oneself). But it has been argued that our knowledge of our own identity is not and cannot be based on any criteria, and hence that this knowledge and the nature of self identity is not affected by the imperfections and inadequacies of the criteria. This point has been made by the argument that the most direct way of knowing personal (self) identity is by remembering\* one's own past history and that our memory of our own past is not, and cannot be, based on any criteria of personal identity. The argument has two steps. The first step is that our knowledge of our own identity (self-identity) is based on or, rather, revealed in, the first-person memory statements that we could make. The truth of this step hardly needs any proof. For it is a necessary truth that if the memory statement is genuine then the rememberer is the same person as the earlier person who witnessed what is remembered. But even when the statement is not genuine, it does not affect the purport of this step of the argument. For in that case, either the memory statement is sincerely made or it is not. If it is not, then the person making it knows that he is not the same as the one who witnessed the thing "remembered"; but if it is a sincere memory statement then as long as it has not been shown to him that he is mistaken, the rememberer is in no doubt about his identity with the earlier person. The fact is that our sincere memory statements carry with them the conviction of our own identity; this conviction is not, and cannot, be shaken by the fact that some of our

---

\*Since what is at issue here is how one knows one's own identity through time, 'remembering' here should be taken in the standard sense in which it is a necessary or conceptual truth that if a person remembers an event then he must have been a witness to that event or must have had direct knowledge of it, whence it follows that remembering one's own past is reassuring of one's own identity.

memory statements may be, and sometimes are, mistaken, just as the fact that our senses sometimes deceive us is no reason that they are never to be trusted. Indeed, Shoemaker has argued that it is a necessary fact that our sincere memory-statements are generally true and that generally they are made only when they are true.<sup>15</sup> This is of course compatible with the possibility of any particular memory-statement being mistaken. But having been obtained from the general case (of genuine memory statements), the knowledge of self-identity is reflected in all sincere memory-statements. Indeed, whether or not one truly remembers, it remains logically true that to claim that one remembers doing x is to claim that one is the same person as the person who did x. Now to the second step in the argument. It is that the first-person memory statements are not grounded on criteria. This point has been convincingly made by Shoemaker<sup>16</sup> by considering the way in which first-person memory-statements are made and known. I will state the situation as briefly and faithfully as is required for my present purpose (which is to bring out the distinctive feature of personal identity).

In the first place, when I make a memory statement about my past, this statement is not grounded on the bodily identity criterion of personal identity; nor, from my point of view, is any consideration of my body relevant to the truth of my statement. It is true that others will need to use, and do use, my bodily continuity in order to be satisfied that I am truly remembering and hence that my statement is true; but I, who is in no doubt about the latter, need no such criterion. Shoemaker went on to say, "from my point of view, it seems inessential that the body I have now be the body I had when I took the walk. If I remember going for a walk then I did go for a walk, no matter what my present body was doing at that time".<sup>17</sup> This is certainly too strong a way of putting the matter; for it is

arguable whether I would be as confident that I did go for a walk if I am given reason to believe that my "present body" was in the study at the time when I remember to have gone for a walk, and thus my bodily non-identity may be relevant to showing the falsity of my memory statement. But once again this does not affect the purpose of the argument that when I do make a sincere memory statement and do believe that I remember what I claim to remember, I do not use and do not need to use the bodily identity criterion of personal identity to make sure that I really remember. Cases of my making sincere but false memory statements may be cases of my falsely believing that I am the same person as the earlier person who witnessed the "remembered" thing. But the issue at stake here is not, whether I am the same person but how do I know that I am the same person. And how I know is the same both where I am the same person and where I only think that I am; and this consists in my being able to make sincere memory-statements in the first-person which does not, and need not, use the bodily criterion of personal identity. In the second place, I cannot be said to use any inner criterion to know that my statement is true. For what this supposed inner criterion could be but my memory? If so, do I have to remember that I remember that I went for a walk? But this breeds the further problem of how this second memory is to be evidenced and thus inevitably pushes the problem backwards ad infinitum. The fact, however, is that I use no criterion as I need none to convince me of the truth of my memory statement (and consequently, of my own identity), although, as we saw, others may use some criterion (that of my bodily identity) to show that I am (or I am not) mistaken\*. But as far as I am concerned, the

---

\*This, however, cannot be taken to show that the knowledge of one's own identity can be grounded on criteria. For as I have shown before (see section 1 pp 7-8,) this will not, strictly, be a case of knowing self-identity but like knowing other identity.

truth of my statement is apparently evident to me by the very fact of the statements being made by me. For it is a necessary truth about the first-person memory statements in particular and the first-person psychological statements in general that they are generally true when they are made sincerely or rather that they are generally sincerely made when they are true. Generally (unless, for example, I am lying) I make a memory statement only when it is true that I do remember, much as I generally (unless, for example, I am pretending or play-acting) say that I am in pain only when I am in pain. The possibilities mentioned in the brackets in each case may justify others, as in fact they do, in looking for some criteria; but they certainly do not affect me, as long as I am sincere and confident in making these statements. There will of course be cases of my mis-remembering; but as long as they have not been shown to be so the state of my knowledge in that case is exactly like the state of my knowledge in the genuine cases and as such is non-criterial. Now, if the first-person memory statements are thus not grounded on criteria and if the knowledge of self-identity is based on, or expressed in these statements, then it follows that the knowledge of one's own identity is not grounded on criteria. What 'shows' to a person that he is identical with someone who existed in the past and did such-and-such things is nothing - not at least necessarily anything - that is criterial evidence for his identity with that earlier person, it is the very consciousness of that identity which operates in his memory of his past. Thus it is that self-identity (personal identity of one's own) is known in self-knowledge. And since, as we have seen, the real nature of personal identity is supposed to lie in self-identity, and not in other identity, personal identity has not unreasonably been thought to be very different from the identity of other material things and has posed a special problem. Not unreasonably, because if one is justified in holding that this speciality is true of one's own

identity then one must be prepared to say that this is true of everybody else, for everyone else, like oneself, is a person, a self-knower and can have non-criterial knowledge of his own identity. It is this "privileged access" which every person has to his own identity that has led some philosophers<sup>18</sup> to treat personal identity as "real" and "perfect" identity and the lack of this access had led them to describe the identity of material things as "imperfect" and "a question of words"(but see Appendix 1 now).

In the previous section of this chapter, I tried to show, though on different grounds from the foregoing that personal identity cannot be defined. And presumably, it is largely because of considerations like the foregoing that Reid and Butler were led to declare that the concept of personal identity is undefineable. By this Shoemaker takes them to mean that there are no criteria for personal identity in the above sense and that one's knowledge of one's own identity, expressed in memory judgements, can not be grounded on criteria.<sup>19</sup> While Shoemaker seems to be generally right in so interpreting their view, I do not think that this interpretation gives quite a complete understanding of their intentions. For at least from something that Reid says it seems rather more likely that "undefineable" is the description which he wanted to apply to the notion of identity as such and not - not at least explicitly - to the notion of personal identity alone. This seems clear from the following:

"Identity in general I take to be a relation between a thing which is known to exist at one time, and a thing which is known to have existed in another time ..."

"If you ask for a definition of identity, I confess I can give none, it is too simple a notion to admit of a logical definition"<sup>20</sup> (My emphasis )

It is needless to point out that personal identity is not "identity in general" and that persons are not the only "things"; hence I think a more faithful interpretation of Reid should have been that what is indefinable is "identity in general" and "personal identity", being an instance of identity, rather what Reid calls "perfect identity", is even more reasonably so. Again, if as Shoemaker thinks being non-criterial were Reid's sole, or even a strong, reason for saying that personal identity is indefinable, he would be prepared to say, for the sake of consistency, the same thing about the identity of material things. But as we see, while he says that "identity in general" (which certainly includes the identity of material things) is indefinable, he would undoubtedly deny that the identity of material things is non-criterial as well. It seems therefore that while Shoemaker may be right in attributing to Reid the view that being non-criterial makes personal identity "real" and "perfect" identity he is certainly wrong in thinking that this was Reid's reason for saying that personal identity is indefinable; for if this were his reason then Reid would not have said what he said in the above passage. On the contrary what was his reason is the fact that the notion of identity (and hence of personal identity which is a specis - a special specis - of identity) is "too simple a notion to admit of a logical definition". Simple predicates are indefinable but not necessarily non-criterial ("identity" when applied to material things, and other persons is not). Thus we can say either that Reid did not say what Shoemaker thinks he said, namely that personal identity is indefinable because it is non-criterial or that he reserved the predicate "indefinable" for personal identity alone (in which case Shoemaker would be right in interpreting Reid in the way he did). But since the latter is counter-textual, and hence false, it follows that Shoemaker was wrong (partly though) in his interpretation of Reid.

I have said in section (i) that the problem of personal identity is not one of definition but one of criteria. (Indicently my sympathy with the view that personal identity is non-criterial in the above sense need not oblige me to give up my original position nor upset it in any way. For in that sense only my own identity as known by me is non-criterial, but not the identity of other persons. But hardly is the problem a problem about my own identity - though, as we have seen, the nature of this identity helps me in understanding the nature and distinctive feature of personal identity. But why it helps me is the fact that as long as I am capable of remembering I can never doubt my own identity - the notion of this is "fixed and precise" (Reid) for me. But the problem which I face and hold to be a problem about criteria is the problem of the identity of other persons. And this is not solved by discovering or being told that self-identity is non-criterial). Perhaps this is also true of the problem of the identity of material things for as we have just seen, it is plausible to say that the identity of the latter is also indefinable. But the difference I want to maintain lies in the fact that in an important sense the criteria of personal identity are more central to, and constitutive of, the concept of a person, whereas this is not true of the relation of the criteria for the identity of material things and the concept of the relevant material things. This I hope to bring out in the remaining part of this chapter. And once this is done this will make more plausible the claim that identity as applied to persons and to material things are not only applied to different subjects but also in different senses - although, I hope, the grounds for this claim has by now been made fairly apparent.

This claim was most explicitly expressed by Butler's distinction between what he calls the "loose and popular sense" and the "strict and philosophical sense" of identity.<sup>21</sup> He thinks that identity when applied to material things is applied in the former sense, and when applied to the case of persons it is applied in the latter sense. His position

... can be roughly explained thus: a tree at present is called the same tree as the plant "it" was fifty years ago. But in the course of this long lapse of time 'it' has undergone serious changes of its parts and elements that compose it; and possibly there is not a single particle of matter in it now which was there in the plant. "And if they have not one common particle of matter they cannot be the same tree in the proper philosophic sense of the word same".<sup>22</sup> If we call it the same tree,<sup>as</sup> we certainly do, this is the loose and popular sense of the word "same."\*

It may seem probable from the above, particularly from Butler's remark just quoted, that identity is ascribed to material things only in a "loose and popular" sense because it is ascribed to them in spite of constant changes that affect them. If just this was Butler's reason for making the distinction between the two senses of identity then he was certainly mistaken. For we have already seen in the last section that identity is

---

\*It may be said (I owe this point to Professor Swinburne) that this distinction which Butler makes is a distinction between two different criteria of sameness rather than between two senses of the word "same". This is quite true, for certainly when we apply the word "same" to material objects and to persons we do not mean two things by that word. The difference in both cases is the difference of what conditions are to be satisfied in order for the word "same" to apply. And this, in effect, is the difference between two sorts of criteria for the application of "same" in each case. However, though putting the distinction in this way is less misleading than the way in which Butler put it, yet the purpose and outcome of Butler's distinction is substantially the same. For we shall see that his distinction aimed at and established the fact that there are two different kinds of criteria of identity as applied to the two different cases.

also ascribed to persons despite such obvious changes. We have seen that this problem of how to ascribe sameness to things which are subject to change is a pseudo problem being based on the mistaken assumption that the ideas of sameness and change are incompatible. This very likely fact that this mistaken assumption might have largely inspired Butler to make the distinction seems to have escaped the notice of Chisholm who saw "an element of truth" in Butler's distinction.<sup>23</sup> However the assumption is mistaken and the distinction would be ill-founded in as much as it was based on this assumption. Therefore the "element of truth" has to be found out in some other, more serious factor. For I think that the distinction Butler pointed to is substantially correct though the reason he may seem to have given for this distinction is mistaken. As Chisholm saw it and as Butler expressed it, the distinction is due to the fact that the identity ascribed to material things is often a matter of words (Butler) or a matter only of decision (Chisholm) whereas personal identity can never be a matter of decision, and any attempt to make it a matter of decision in any otherwise compelling circumstances can succeed only at the cost of making a significant change in the concept of a person.

This way of putting the distinction is less misguided and more helpful in understanding the peculiarity of personal identity. We have just suggested that material objects' constantly changing is no reason for saying that their identity is to be described as "loose and popular", nor, for that reason, will it be a matter of decision. For if this were the reason then even personal identity would not escape this charge (See sec.1 )

What then is the reason for the distinction and what is the reason for our thinking that the distinction is substantially correct? To this purpose let us travel along the path which Chisholm takes since, I hope, this would show us the reason.

Chisholm explains this distinction by means of his example of the U.S.S. South Dakota - which is but a different version of the well-known case of the "Ship of Theseus". It is supposed to be a ship whose wooden planks are replaced, one at a time, by aluminium ones so that one day it becomes totally made of aluminium. We would still call it the same ship - the U.S.S. South Dakota. But we are also asked to imagine that all the original wooden planks, without any alteration or damage to them, are simultaneously being reassembled into another ship. Now the question is which of the two ships, the wooden one or the aluminium one, is the U.S.S. South Dakota? For the 'continuity of form' criterion demands that it is the aluminium one, while the 'identity of parts' criterion demands that it should be the wooden one, because "after all, it is made up of the very same parts, standing in the very same relations ...".<sup>24</sup> There is, strictly speaking, nothing<sup>25</sup> to choose between the two criteria of identity here (see note 25 now). Hence, Chisholm argues, the question is not; which of the two is the same ship?, but, which of the two is to be counted as the same ship?, and this is a question to be decided by a court of law or by the appropriate authority by deciding upon an answer. But although it would not be very unnatural whatever decision is taken, yet any decision they take would be fairly "defeasible" and can be quite naturally challenged and superseded by another decision by a superior court or authority. By contrast, when we say that a later person is the same as an earlier person, Chisholm says we do not do it by any decision. In this sense, he says, the criteria of ship-identity is "a matter only of convention" and this is not true of personal identity. But it may well

be asked, why can't this be true of personal identity? Let us imagine a somewhat analogous change to happen to a person. Suppose a person A, is found one day to have undergone serious and total change of memory, character and personality. The A-body person (the person who now speaks from the body that, prior to the change, was A's) is found to possess all the opposite characteristics from what A used to have. He is found further to make memory claims which, when checked, do not at all tally with what A had been doing so far. Is this person the same person as A or a different person? We will not know what to say; the 'continuity of memory' criterion demands that he must be a different person, whereas the 'bodily identity' criterion demands that he is not a different person. And here again there seems to be nothing to choose between the two criteria. But we will have to say something - whether or not he is the same person as the earlier. And whatever we say, will it not amount to our adopting, without sufficient reason, one criterion to the total exclusion of the other? Hence, will it not then be a matter of decision and does it not make the criterion of personal identity equally a matter only of convention? The defence of Chisholm's view comes not from Chisholm himself - not from him explicitly - but from Shoemaker in his "comments" on Chisholm's paper.<sup>26</sup> After attacking Chisholm on the same score - though on a slightly different ground,<sup>27</sup> Shoemaker points out that when we adopt a convention in calling one of these ships the same as the original, we do not change the meaning of "ship" significantly. As he rightly argues, "it would not be natural to say that British judges mean something different by 'ship' from what the American judges mean simply because, what is imaginable, British Courts rule that the wooden ship is the U.S.S. South Dakota while the American Courts rule that the Aluminium ship is the U.S.S. South Dakota."<sup>28</sup> And he substantially agrees with Chisholm's intentions that there would be a significant change in

the meaning of "person" if we adopt the convention of calling him the same person only by virtue of his having the same body or, alternatively, if we decide to say that he is a different person only by virtue of his having different memories and character, for in the former case what, in effect, we would be judging to be the same is a continuing body - which is tantamount to saying that a person is a mere body; and in the latter case what we would be judging to be different (non-identical) is something not - at least not essentially - bodily since it is judged to be different despite the observed spatio-temporal continuity of the body - and this is tantamount to saying that a person is something not essentially bodily.\* Thus, by taking the decision one way or the other we would be subscribing either to a materialistic concept of a person or to a cartesian concept. In either case we would be making a significant change in the meaning of "person". For the normal concept of a person is that to which both a set of M-predicates and a set of P-predicates (of Strawson) are ascribed. We get to know the meaning of "person" by being shown things (or beings) who have both the characteristics, who shake hands with us and greet us - by being told not only that they are tall or short, fat or slim, but also that they are intelligent or dull, thoughtful or unthinking. These two sorts of characteristics - physical and psychological - are each essential but neither is sufficient to the concept of a person. This fact about the nature of the concept of a person led Shoemaker rightly to say that "the concept of a person is to a considerable extent defined or constituted by the criteria of personal identity .... and any attempt

---

\*The materialists and the Cartesian mentalists, who obviously mean different by the concept of a person, tend unsurprisingly to adhere to opposing criteria of personal identity.

to make a substantial change in the criteria (by deciding to adopt one criterion to the total exclusion of the other, as we saw, PKM) is likely to succeed only in changing the meaning of the word person i.e. making it express a different concept from what it does now". "The suggestion is that the criteria of personal identity are much more central to the concept of a person than are the criteria of ship-identity to the concept of a ship".<sup>29</sup>

The suggestion seems reasonable enough and the ground on which the suggestion is based seems equally reasonable. It is significant that this ground is peculiar to the case of persons and personal identity and that it cannot be generalised. In other words, not every concept is determined by the criteria for its identity. For not every word would express a different concept if we adopt a different criterion of identity. For as suggested by Chisholm and shown by Shoemaker we will mean little different by "ship" if we adopt the one criterion or the other of "same ship" by taking as we must in the case described, a decision. This is largely true of material things. But we saw by analysing the concept of a person that a decision in favour of one criterion to the exclusion of the other would make a significant change in the concept of a person. Wiggins said "it is the sole peculiarity of persons that it is more than usually odd to call it a decision of ours to employ this sort of

30

The oddity is due not only to the fact, just noted, that the criteria of personal identity are more central to the concept of a person, but also to the fact noted earlier, that persons are self-knowers. Let us see how. With a view to making Chisholm's case plausible, I imagined a case where a person undergoes serious changes. In that case, I tried to show that since there would seem to be nothing to choose between the two competing criteria, of bodily continuity and continuity of memory and character, each pointing in opposite directions, application (or otherwise) of

sameness in that case would appear to be equally a matter of decision. But the appearance is deceptive, for although we may be persuaded to think that taking a decision is the only resort yet there is a sense in which we could be wrong about our decision. For there would be at least one person who can supply a good reason to prove us wrong; it is the person himself whose identity is in question. Since he is a person and hence a self-knower and can have the most direct (non-criterial) knowledge of who he is, our decisions are of no use to him. For him the question of his identity is not "a question of words" in any sense.<sup>31</sup> On the contrary whatever decision we take it has always a danger of being wrong by a shrug of his shoulder. But once again, if this is true of him because he is a person and hence a self-knower, it is also true of all of us as we are also persons and each of us has this "privileged access" to our own selves (But see Appendix 1). Thus it is because of the concept of a person, the fact that to be persons is to be self-knowers, that it is odd to say that the criteria of personal identity is a matter of decision. This also adds to the ground for saying that the concept of a person is more intimately connected to the criteria of personal identity.

Of course by claiming that persons are self-knowers it is not denied that people could be sometimes unsure of, even wrong about, their own identity, but what is important is that they can be, and generally are, right about it. It is this fact that people can make identity-judgments about themselves and are in a position to make them without using any criteria and without needing any justification of their judgments, and further that they can be, and generally are, right about them, which constitutes an important difference between the identity of persons and that of other material things, and accounts for there being a special problem about the former.

PART - II

## Chapter 2

### THE TRADITIONAL APPROACH

#### Prologue

So far I have exposed the nature of the problem of personal identity and proposed to discuss the problem in its aspect of reidentity or diachronous identity of persons. An important outcome of our discussion of the problem - especially in the first section of the previous chapter - has been that the problem involves two distinct questions. To owe a formulation of Professor Swinburne, (See his 'Personal identity' PAS 2974), the first question is: what does it mean to say that a person  $P_2$  at time  $t_2$  is the same person as the person  $P_1$  at time  $t_1$ ? (I shall refer to this as the "meaning - question"); and the second question is: what evidence can we have that a person  $P_2$  at time  $t_2$  is the same person as the person  $P_1$  at time  $t_1$ ? (to be referred to as the "criterion - question" \*). Anyone asking the first question will be satisfied only by being provided with a definition of personal identity. But someone asking the second question need not look for a strict definition, and will be satisfied with the criteria for determining the sameness of persons at different times. An answer to the first question - if one could be found at all - would, of course, answer the second, since the defining feature(s) of  $x$  will provide the most satisfactory criterion (criteria) of  $x$  - occurrence; the  $x$  reverse, though, is not true. But, as I have shown in the previous chapter, no satisfactory non-trivial definition of personal identity can be possible and, for this reason, it has been suggested, any attempt at defining

---

\* For what is required, I maintain, is not just any evidence, but only what are the criterial evidences of personal identity.

personal identity will not be on the right lines towards a solution of the problem at stake. I have therefore emphasised that our problem properly is one about the ~~the~~ criteria of personal identity; and while the problem must be thus understood the specification of the requisite criteria need not, and will not, be affected by the unavailability or impossibility, of a definition of personal identity.

Now, although the two questions noted above are thus distinct and the understanding of their distinction thus important, this has not been often realised. As I shall show in this chapter, most traditional theorists (some contemporary writers not excluded) have not merely made no clear distinction between the two questions but, taking the problem as one of meaning, some of them have tried to define personal identity in terms of what can only be a criterion thereof. In fact, it seems to me that this feature ( of non-distinction of the "meaning-question" and the "criterion-question " and taking the problem in the sense of the former) constitutes a distinctive mark ~~to~~ of the problem. I cannot pretend, of course, that this feature is entirely absent in contemporary writings on the subject, but I do hope to bring out in the subsequent chapters that the tendency there is largely to keep the questions apart, though I shall also contend that this tendency did not obtain full possession of the minds of all contemporary writers. Meanwhile my concern in the present chapter will be to analyse some important traditional answers to bring out the described weakness in them. And I shall show in the process that taking the problem in the sense of the "meaning - question" <sup>are</sup> inevitably leads to solutions that ~~are~~ either mystical or otherwise strange. My analysis of these theories will also help to unearth the basic philosophical assumptions of the respective philosophers that led to the peculiar solutions they gave to the problems.

Section (i) Locke and Paine: the memory theorists

The first philosopher ever to have realised the importance of the problem of personal identity and to have given a systematic account of it was John Locke. His treatment of the problem occupies the 27th chapter of Book II of his Essay. He seems to have thought a theory of personal identity to be of importance for the purpose of moral accountability. For him person is a "forensic" term (26)\* and personal identity is the foundation of "all the rights and justice of reward and punishment"(18). It is of course obvious to any theory of justice and moral accountability that no person should be punished (or rewarded) for what another person did and that reward and punishment should go to the person who did the action in question. Some even have claimed that it is logically impossible to punish someone for something he did not do - though you may inflict pain in him<sup>1</sup>. And so it is imperative on such theories to establish that the person who is going to be punished or rewarded for a certain action is the same person who did the action. But although something like a "forensic" consideration might have made Locke realise the importance of the problem of personal identity, he seems to have gone far beyond the problem of establishing who is who to the much deeper problem of the meaning of personal identity. His concern seems to be not simply, what evidences we can have (or how do we know) that a later person is the same person as an earlier person, but, what makes the later person the same person as the

---

\*Bracketed numerals throughout this section refer to the sections in Chapter 27 of Bk II of Locke's Essay.

earlier person. The identity of a person may be established on evidences; but it would be absurd to say that these evidences make a person the same person as an earlier person or that they constitute the meaning of "same person" (our discussions in the first section of the last chapter has made this considerably clear; and more will be said on this in Ch 4 sec (1).) And whereas the availability of relevant evidences may solve the "forensic" problem, it cannot solve the problem of the meaning of personal identity.

That Locke was concerned with the meaning of personal identity would be fairly clear from a careful reading of his chapter on "Identity and Diversity"<sup>2</sup>. It was noted in the last chapter (section 1) that "identity" is an incomplete term which needs to be conjoined with a substantive in order to be applied. And what would be the nature of identity, and how it would be determined, depends largely on the nature of the substantive concept to which it is applied. Once again, Locke was first to realise this vital point and to insist that "such as is the idea .... such must be the identity" (see esp (7)). Accordingly, and very justly, Locke's account of personal identity was based on his notion of a person. By "a person" he means "a thinking, intelligent being that has reason and reflection and can consider itself as itself, on the same thinking thing, in different times and places".<sup>3</sup> The first part of this quotation (not emphasised) constitutes Locke's definition of a person and the rest (under emphasis) gives his notion of personal identity, which he in other words describes as the "sameness of a rational being". A rational being can consider "itself as itself in different times and places" - can know of its identity through time, says Locke, "only by means of that consciousness which is inseparable from thinking and essential to it"(9). And not only that a person knows his identity by means of what Locke calls "consciousness" but also that "in this (consciousness, PKM) alone consists personal

identity" (9, emphasis mine). It seems clear from what has just been said above that Locke, like Reid and Butler (vide Chapter 1 sec ii), takes self-identity as the paradigm case of personal identity, but unlike them, he takes "consciousness" as constituting personal identity. Unfortunately, however, Locke's use of the word "consciousness" is not unambiguous, and there are at least two different senses in which it can be taken. At some places, Locke used "consciousness" in such a way that it may easily be taken as "self-consciousness", e.g. "a being... that can consider itself as itself .... does so by that consciousness which is inseparable from thinking." (my emphasis). This is how Flew quoted<sup>4</sup> Locke to show the same point; and thus quoted Locke certainly seems to have used "consciousness" in the sense of self-consciousness. But here a phrase is missing from this quotation which I think is very significant for the problem of personal identity, which I have taken to be that of reidentification (see the abstract); and that missing phrase is: "in different times". With this phrase inserted, as it originally was, after the phrase under emphasis, "consciousness" would mean - partly at least - memory. For what sort of consciousness it is - if not remembering one's earlier selves - by which one considers oneself as oneself at different times ? However, I will not dispute the claim that the above sense (of self-consciousness) can be read into Locke's use of 'consciousness' at some places; as a matter of fact there are clearer evidences (and happier passages) to show that e.g. "consciousness always accompanies thinking, and it is that that makes everyone to be what he calls 'self' and thereby distinguishing himself from all other thinking beings". (9). But what seems to me to be more plausible is that when taken in this sense "consciousness" can serve Locke at best as giving a part of the meaning of "persons", but can never be of any help or relevance to that of personal identity - unless it involves or means "memory" in some way. And I suggest

that this was Locke's intention when he claimed that consciousness makes personal identity. This is evident from: "as far as this consciousness can be extended backwards to any past action or thought so far reaches the identity of that person" (9, emphasis mine), and from "it (consciousness, PKM) is but a present representation of a past action" (23) and more explicitly from "could we suppose any spirit wholly stripped of all its memory or consciousness of past actions ....."(25, emphasis mine). But we have seen that memory, as it concerns personal identity, is only a criterion of personal identity and could not serve as (even) a necessary conditions for it. Yet to Locke, it was not merely something that established the accross - time identity of persons but something that entails that identity - something in which personal identity consists. In his famous prince/cobbler case (15), Locke says, for example, "should the soul of a prince carrying with it the consciousness of the prince's past action, enter and inform the body of a cobbler..... he would be the same person with the prince"(emphasis added). Thus it is that, for Locke, what is important in making the cobbler - body - person the same as the prince is the memory ("consciousness of the prince's past life" ) of the prince - his soul being not enough, presumably because Locke was not quite convinced about its nature and existence. (cf 10) Locke followed this principle to whatever strange consequences it led him to go. He was, for example, prepared to say that "if Socretes and the present Mayor of Queensborough agree ( in respect of consciousness, PKM), they are the same person" and even that "if the same Socretes waking and sleeping do not partake the same consciousness, Socretes waking and sleeping is not the same person" and what is still worse, "to punish Socretes waking for what sleeping Socretes thought, and waking Socretes was never conscious of, would be no more of right than to punish

one twin for what the brother twin did" (19) These consequences are strange ( and Locke himself is found to be struck by the strangeness of his consequences), but inevitable if one is to hold, like Locke did, that the identity of a person must follow the destination of his memory and other mental characteristics.

We noted in the beginning that Locke emphasised the importance of a theory of personal identity so that rewards and punishments may go to the right persons; presumably this was one of the motives for Locke's analysing personal identity in terms of consciousness or memory. For usually, consciousness of having done the action is an important ground for ascribing responsibility to a person - thereby implying that this person is the same person as the one who committed the crime. But (what seems to be) Locke's mistake was to overlook the fact that this is neither the only means of establishing the identity of a person nor is the only ground for ascribing responsibility. On the contrary, we are often prepared to take up responsibility for many of our forgotten deeds, if we could be made to believe that we did do these. And where consciousness of guilt is the sole ground for ascri<sup>b</sup>bing responsibility it does not follow that it makes the same person or that the lack of it makes a person numerically different. Locke, who seems to be convinced that consciousness makes the same person (and conversely), invokes the example of human laws which do not punish the mad man for the sober man's actions and the sober man for what the mad man did (20) to argue that these laws (and by implication, lack of consciousness or memory) make them two different persons. But this is utterly absurd. For in the first place these laws can at best be said to treat them as two persons, but not to make them two persons. Secondly, we do not mean two different things by the pronoun "the" when we say that he committed the crime, but

that he should not be punished now since he is completely mad. And it can never be disputed that our reference here is to the one and the same person who committed the crime and who, on extraneous grounds, is pleaded to be exonerated. Moreover, what one can plausibly argue here is that what this example proves - if it does prove anything - is that it is not the person as such but the person with a particular personality and character that is the recipient of rewards or punishments. And pace Locke, instead of "person" being a forensic concept, "personality" and "character" should be more plausibly so.

The weakness of Locke's theory that consciousness makes personal identity became conveniently obvious to post-Lockean critics. First there is the famous objection of Butler<sup>5</sup> that consciousness presupposes personal identity and therefore cannot be said to constitute it, without begging the question. To say that one is conscious of some past action (that is what we have seen Locke's use of "consciousness" to mean as relevant to personal identity) is to presuppose that one is the person who both did the action and who is now remembering the doing of it. But it is this very identity of oneself with the person who did the action that is in question, and which consciousness is supposed to constitute: If consciousness or memory is taken in the strong sense, there seems to be no escape from Butler's objection. If, on the other hand, it is taken in the weak sense of, apparent memory (viz, honest but false memory and which currently is described as q-memory\*), the charge may be avoided; but then Locke's definition of personal identity would run into the difficulty of allowing not only that the victim of what is called paramnesia is the

---

\* When it does not mean what is called genuine q-memory. But this last, being q-remembering one's own experiences would be indistinguishable from memory in the strong sense and hence cannot escape Butler's objection.

same person whose actions he putatively remembers, but also that two (or more) contemporaneous persons claiming to remember what a third, long since dead, person did would be identical with the latter without, for well known reasons, being identical with each other.<sup>6</sup> Locke is seen to be perplexed by the puzzling consequences issuing from the extreme openness of "memory". Yet unwilling to give up memory as the defining feature of personal identity he was reduced to appealing to the goodness of God to prevent such untowardness from occurring (¶ of 23 ). Secondly, besides difficulties of this type there is the other difficulty arising out of the possibility of amnesia or loss of memory. This is expressed by the well known case of the gallant officer<sup>7</sup> who was flogged at school, took the enemy's standard as a young officer and was later on made a general. It is conceivable that the officer who took the standard remembered, at the time of taking the standard, having been flogged, but that the general did not remember the flogging though he did remember having taken the enemy's standard. By the logic of transitivity of identity, if the officer who took the standard is the same person as the school boy and if the general is the same person as the young officer (each pair must be identical on Locke's definition), then the general must be the same person as the school boy. Yet, on Locke's very same definition, the general cannot be the same person as the boy since his consciousness is not "extended backwards" to what had happened to the boy. This is another inescapable predicament<sup>8</sup> that faces a theory which tries to analyse personal identity in terms of memory or what Locke called the continuity of consciousness (but see note 8).

It is interesting to note that being eager, as he was, to define personal identity Locke defined it in terms of memory which is but one of the criteria of personal identity: the criterion of bodily identity and

even the quite general fact that persons have bodies were treated by him at best as contingent matters of fact. As the foregoing considerations would explain, being persuaded to think of persons as essentially non-physical things, Locke remained a memory theorist in his account of personal identity - the latter being the logical outcome of his concept of a person.

Hume, likewise, remains a memory theorist, though not on the same ground. For he did not think that there is anything really to be defined, or even described, as personal identity. After having reasoned that everything that lasts beyond a moment is only a "bundle or collection" of rapidly changing perceptions he concluded that the identity that we ascribe to persons ( and to other things) is only "a fictitious one".<sup>9</sup> That is, he was led to this position on the ground that there can be nothing permanent in a person to account for his unity and identity. This is brought out by attacking the notion of self. Hume argues that if every idea has to come from an impression then there can be no idea of the self (and hence of person). For, in the first place, there can be no impression for the idea of the self to be derived from; since it is that to which all our impressions are supposed to refer. Secondly, there could be no such impression which is constant and invariable, and yet "the self is supposed to exist after that manner"<sup>10</sup> From this Hume concludes that there can be no permanent self to continue the same so as to account for the unity of the mind.

Now there are two facets of Hume's argument which I will consider before proceeding further. For it seems to me that in both cases the premiss does not warrant its purported conclusion. The two facets of the argument are:

- (1) Our impressions of objects are different at different moments;

so no object of one moment can be the same as an object at another moment.

- (2) Every idea is derived from some impression; so as is the nature of the impression, so must be the nature of the object of which the idea is an idea (whence it was supposed to follow that, since no impression can be constant and invariable, nothing can be constant or permanent).

As regards (1), it can be said that my impressions of a table at different moments may be different; but it would be absurd on this ground to say that the table, which my impressions and idea are 'of', is a different table at those different moments. And regarding (2), it can be said that there cannot be any impression, and hence any idea, of the whole world in the sense that no one can form such an impression; but from this it does not follow that there is no such thing as the whole world. The error in these ways of reasoning lies in the fact that it is generally wrong to deduce anything about the nature of things from the nature of the impressions from which the ideas of the objects may be derived. The impressions which I have of my writing table may not be continuous and may be interrupted by some other impressions and ideas. But does it make my writing table discontinuous and numerically different? Yet a psychology somewhat of this sort seems to be at work behind Hume's theory.

Another characteristic element of Hume's thought, which makes him a memory theorist, was his enormous reliance on the application of the "inner test". "When I enter most intimately into what I call myself I always stumble on some perception or other .. I never can catch myself at any time without a perception and never can observe anything but perceptions".<sup>11</sup> As the inner test never revealed to him anything permanent but

only the fleeting perceptions in him, he concluded that there was nothing permanent in him. This in turn led him to describe man (he meant nothing different by "man" and "person") as a "bundle of different perceptions". And what puzzled him was the fact that our understanding never perceives any real connection between these different perceptions. Consequently, he argued that identity is nothing really belonging to these perceptions and uniting them together into a single self or person, but that it is merely a quality we attribute to them. If however, we are to talk of "same person", presumably for moral and practical purposes, the only criterion we could use is memory. For "as memory alone acquaints us with the continuance and existence of this succession of perceptions it is to be considered as the source of personal identity"<sup>12</sup>.

I have tried to show that the lack of any permanent impression does not show that nothing permanent could be there; and if I am right this would also weaken Hume's dismissal of a permanent self and the consequent denial of personal identity. But some have also argued that Hume's inner test did enable him to observe his self, or else it did not show him anything - not even the fact that he stumbled only upon his perceptions. Chisholm<sup>13</sup> has argued that when Hume "stumbled" upon different perceptions, it is implied that he found himself stumbling upon these perceptions, and hence that it would be self-stultifying for Hume to say that he did not find himself. Although this sort of consideration, in a sense, guarantees the knowledge of oneself (in the sense in which we have said, see Ch 1 sec ii, that persons are self-knowers), yet one feels uneasy by Chisholm's purported claim that when we know, e.g. that we are thinking, we know our selves to be thinking. Whatever persuasive force this argument may seem to have, it is utterly mistaken. It looks persuasive because no one could certainly deny that when he knows that he is, e.g.

watching cricket he also knows that he is watching cricket. Yet the argument seems to be mistaken for at least two reasons. Firstly, the alleged knowledge-claim of oneself in such cases is entirely vacuous, since what is known here is simply my watching cricket, and it will be very odd indeed to say that in knowing this I know two things - that I know the "I" in me and know that "it" is watching cricket. Secondly, the argument makes a fundamental confusion\* between the technical sense of 'self' in which it means a permanent substantive, immaterial something (which, allegedly, accounts for the unity and across-time identity of persons), and the ordinary sense of the word in which it is usually prefixed by possessive adjectives (like myself, yourself etc) and in which it does not have any such implication as the above. When Hume or anyone like him denies the knowledge of himself, he is using "self" in the former sense and this denial is perfectly compatible with his assertion that he knows that he is (or that he knows himself to be) watching cricket or "stumbling on some perception", since the "self" in this assertion is used in the ordinary non-technical sense; and familiarity and naturalness of these ordinary assertions will not affect Hume's arguments against the existence of the self or soul.

However, I do not think that the knowledge of the self in the technical sense is necessary to explain personal identity, nor will its denial entail the unreality of personal identity. For if the self was thought to meet the requirement of "something unchanged" to account for the unity and identity of persons, then we have already seen (see Ch 1 sec 1) that it is a spurious concept; because the idea which generated a false need for such a concept - namely the idea that sameness and change are incompatible - has been shown to be mistaken. And in so far as Hume fell victim to this mistaken idea he was wrong and his denial of real (as opposed

---

\* The spirit of this argument was suggested to me by Professor Swinburne.

to 'fictitious) identity to persons, based as it was on this mistaken idea was misguided.

However, as we have seen, despite his theoretical position, Hume was willing to consider, for moral and practical purposes, the nature of personal identity. But whatever he had to say on this was vitiated by his insistence on the use of the 'inner criteria'. This has led him to give a very unrealistic account of personal identity; for he is seen to ask and consider the question only about his own identity and never seems to be concerned with the identity of others besides himself - though it would hardly be disputed that the problem, as encountered in real life, is generally about the latter.\* This limitation of Hume had two consequences which contributed towards the weakness of his theory of personal identity. First this limitation made Hume's account of personal identity uninteresting. Secondly, it was at least one of the reasons why he totally ignored the bodily considerations in his account of personal identity and became some-what of a dogmatic memory theorist. To take the second point first, while considering the question of one's own identity nothing necessarily bodily comes to the picture and all one considers is what actions one remembers to have performed or what experiences one remembers to have had in the past. Since memory thus gives us the sense of our identity, bodily considerations tend to be ignored and seem to be inessential. But this is a gross over-simplification. For when you are considering whether you are the same person who did such-and-such things yesterday, you are not only concerned with your memory and inner perception, you are also referring to your doing such-and-such, and this has an essential reference to your body which therefore has a role to play -

---

\* This has been argued fully in Ch 1 sec 1 (see pp 7 - 8 above)

however unimportant you may prove the last to be. Further, there may be possible situations when one may doubt who one is, and in such cases one has to depend on what others say. But what others will say depends, pre-eminently, on their judgments about one's bodily continuity (since one's memory now is put to question, even for oneself) and this is just what is missing in Hume's account of personal identity. Being solely limited to his own identity and its inner tests, Hume is not only unconcerned with other persons and their identity, but is in a way ruling out the logical possibility of their being there at all. For if there are others, Hume, by ignoring the bodily continuity criterion and insisting upon the inner criterion, has no means of identifying them. But then - and this brings us to the first consequence mentioned above - this makes one's own identity pointless. As Pears pointed out<sup>14</sup> self - identity must go side by side with other-identity. Had there been no possibility of there being other persons, the question of one's own identity would not arise, "for there would be nobody for me not to be".<sup>15</sup> Consequently the account of personal identity would be uninteresting.

Section (ii) Butler and Reid: the intuitionists

The Lockean-Humean theory of personal identity had two lessons for posterity. The first, which was a positive lesson, was that it had emphasised the fact that the consideration of memory is immensely important to a theory of personal identity. But the second lesson, a negative or indirect one, was that it revealed - by its failure - the important fact that memory cannot be said to constitute personal identity, that it cannot make a person the same person as an earlier person. In other words, the limitations of their theories made it clear that personal identity cannot be defined in terms of the continuity of memory or consciousness. These factors are reflected in the writings of Bishop Butler and Thomas Reid on this subject.<sup>1</sup> Both these philosophers wrote almost during the same period and both seem to have reacted to the Lockean-Humean theory in similar ways, which perhaps, is why their views on the subject remains largely the same. Both have taken memory as the most reliable evidence (Reid calls it "the most irresistible evidence" Ibid P205) of personal identity and both have claimed that personal identity cannot be defined<sup>2</sup> at all much less in terms of memory. And what is something both seem to be particularly critical about is the theory, mostly due to Hume, that we are, at every different moment, a different person and so cannot be really the same at different times. Although Locke does not explicitly profess such a Humean theory, Butler and Reid argued that he is committed to it by making provisions for it. Butler can be seen to be making this point when he says that a person cannot really be the same at different times if, as the Lockean says, the consciousness in which personal identity consists is not the same individual act of consciousness at different times.<sup>3</sup> And it is not difficult to see that Reid would attribute such a theory to Locke, though in a more indirect way. For Locke allows the possibility that the substance that

thinks in a person may be different at different times,<sup>4</sup> and for the sake of consistency, Locke must admit that a substance and a person are the same in so far as each is defined by him to be "thinking thing"<sup>5</sup>.

Reid's critical attitude towards such a Humean theory, which in effect claims that the identity we ascribe to persons is only a fictitious identity, is given vent to right at the beginning of his account of personal identity. For he starts by saying: "the conviction which every man has of his identity... needs no philosophy to strength it, and no philosophy can weaken it without first producing some degree of insanity".<sup>6</sup> And Butler's similar reactions to the theory in question is expressed in his claim<sup>7</sup> that "this notion is absolutely contradictory to that certain conviction which necessarily and every moment rises within us...." We shall in due course come to see that their reliance on this "conviction" that is supposed to acquaint us with our identity makes them memory theorists with a difference. For although they take memory as the most reliable evidence of personal identity, they hold that the latter is something "simple and unanalyzable" and that its real nature can be known only by means of some sort of non-sensuous intuition.

Butler's desertation on personal identity arose, partly at least, out of his criticism of Locke. Like Locke, he gave consciousness or memory the prime importance in his account of personal identity. He was rather justly of the opinion that consciousness or memory of our past actions and experiences gives us the very sense of our identity through time and assures of our identity - at least with the earlier persons who did the actions and had the experiences in question. For my remembering that I did such and such a thing in the past carries with it the conviction that I am the same person as the earlier person who did that thing, and in an important presupposes this identity. This logical connection, which memory is said

to have with personal identity, is an undeniable fact in all standard cases where memory is taken in the sense of "real" memory, and this, if anything, is the rationale of our inclination to use memory as a criterion of personal identity. But if this is so, Butler argued, then it would be absurd to claim that personal identity consists in memory or can be defined in terms of it. In its familiar formulation, Butler's objection is expressed by saying that consciousness, or memory in the relevant sense, presupposes personal identity and therefore cannot constitute it. We have seen (sec i) the effect of this criticism on Locke's theory and something more will be said on this aspect of the criticism later on. What now needs mention is the fact that Butler's objection has another aspect which has been given rather insufficient attention than what it deserves - though I think this is more effective against Locke than the other, familiar, point of the objection. To bring this out, I will quote the objection fully:

"But though consciousness of what is past does thus ascertain our personal identity to ourselves, yet to say that it makes personal identity, or is necessary to our being the same persons, is to say that a person has not existed a single moment, nor done one action but what he can remember; indeed none but what he can reflect upon. And one should really think it self-evident, that consciousness of personal identity presupposes and therefore cannot constitute personal identity; any more than knowledge, in any other case, can constitute truth which it presupposes".

(Analogy p329, my emphasis)

It is easily seen that Butler's first point as expressed in the first sentence above, makes the Lockean theory more palpably absurd. For we forget many things and many memories of our own doings are buried in eternal forgetfulness; and on Locke's theory we will not be the same persons as the persons who did these latter things. If the definiens is absent the

definiendum must be absent as well. Further, this criticism of Butler shows not only the impossibility of defining personal identity in terms of memory, but even of the latter being a necessary condition of the former.

Now, back to the familiar point of Butler's objection. We have already seen in the previous section that Locke has no easy escape from this. However the efficacy of this objection has been questioned by some recently. Palma,<sup>8</sup> for example, complained, first, that Butler's use of "consciousness", particularly his use of "consciousness of personal identity", is ambiguous, and secondly, that if it is taken, as it usually has been, in the sense of memory then it fails to bite, since there cannot be any "memory" of personal identity in the strict sense. I shall say that the first point in this criticism, though not entirely without a ground, is trivial and ill-founded. For although Butler's use of the phrase "consciousness of personal identity" is slightly odd, and misleading, his use of "consciousness of what is past" in the preceding sentence and elsewhere should be enough to make that up and to make it clear that by "consciousness" he means memory. Moreover, we have seen while considering Locke that the only kind of consciousness that is relevant to an account of personal identity is that kind which is memory - the other kinds of consciousness, like self-consciousness (where it is not memory) or consciousness of something present, though of peripheral interest, are not as relevant as memory. Palma's second point lays down a sound philosophical principle; but the principle has now use in the case in question. The principle is that nothing present can be properly said to be "remembered",<sup>9</sup> and hence that personal identity, which is a present attribution on the basis of the present representation of somethings past, cannot be remembered. It is certainly absurd to say that I remember that I (myself at the present moment) am the same person who did such-and-such things. But never is

Butler committed to saying this. All he would be claiming is that if I remember doing such-and-such then I am the same person who did such-and-such. The "consciousness" which Butler claims to be presupposing the identity of the subject is this consciousness that he, the subject, did such and such. This is the spirit of Butler's objection which, I think, is fairly clear from what he says - at least in the passage quoted above. The trouble with what I call the Palma-type\* critics seems to be that they want to insist that if at all Locke and Butler (or any one of them) can be said to mean memory by "consciousness" then they can be said to mean nothing else elsewhere by that term. But this is unduly to overlook the important fact that consciousness has various forms and that a fair account of Locke's and Butler's views should be that only that form of consciousness which is called memory is relevant to, and is said to have the relevant logical connection with, personal identity. Saying this does not commit Butler and Locke to use consciousness only in the sense of memory and not in any other sense anywhere else. As we have seen earlier in the last section (see 1990-Galov, "consciousness", even as Locke used it, is not ambiguous between memory, self-consciousness, consciousness of something present and so on; it signifies a generic psychological state or function of which the latter are specific forms.

Now, on the constructive side of Butler's theory of personal identity, there is the well-known distinction between the two sense of

---

\*By this I mean roughly those who claimed that Locke did not, and Butler cannot, mean memory by 'consciousness' and hence, by implication, that Locke was not a memory-theorist and that Butler could not use his objection without absurdity.

identity. Like Locke and unlike Hume,\* Butler believed that "same", as applied to persons and to other things, is applied not only to different subjects but also in different senses. The sense in which sameness is ascribed to persons in what Butler called the "strict and philosophic" sense and the sense in which other things are called the same is what he described as the "loose and popular" sense. We have noted before (see Ch1, sec ii above) that it is rather misleading to say that the distinction is between two different senses of identity and that what this distinction should, and does, point to is the fact that there are two different criteria for the application of the word "same" or "identical" in the two different contexts. We also observed that although Butler's distinction is substantially correct, yet the ground on which he seems to have made this distinction is mistaken. For in saying that the identity of the ordinary material things, like trees and ships, is identity in the "loose and popular" sense, Butler was surely reasoning like Hume and was thinking that whatever is subject to change cannot really be the same. He gives yet another proof of his falling prey to this dogma when he says "But in a strict and philosophical manner of speech, no man, no mode of being, no anything can be the same with which it has indeed nothing the same."<sup>10</sup> I have nothing more to add here as regards the significance of the distinction Butler made than what I have already said in Ch i sec ii, namely that while the identity of the ordinary material things can be, and often are, a matter of decision the identity of persons cannot, without

---

\*Hume too, it may be recalled, made a distinction between two sense of identity, i.e. between what he called "fictitious" identity and "real" identity. But this, I will say, is little to the purpose. For one thing, even though Hume made this distinction he nevertheless claimed that identity is applied, if applied at all, in the same sense, i.e. of being fictitious, when it is applied to persons as well as to other things (see Ch 1 sec ii above). Secondly, if Hume was right, his sense of 'real' identity has no application. For wherefrom did he derive this idea, if not from any impression? And if, on Hume's own principle, all ideas are to be derived from some impressions, then this idea of 'real' identity is unreal and so his distinction between 'real' and 'fictitious' identity cannot work.

producing some degree of absurdity, and without making some significant change in the concept of a person, be said to be a matter of decision, and that while there can be wrong or right answers with regard to the questions of personal identity there is often no such things in the other case - in some of these cases we can say whatever we like or whatever that suits certain pragmatic considerations, without making any significant change in the concept of the things in question (see pp     above). But while our reason for saying this is the peculiarity of the use of the concept "person", Butler's reason seems to have come from the idea of something "intrinsic" in the nature of persons (see the strict sense). "cannot subsist with diversity of substance".<sup>11</sup> By this I take him to mean that despite the changes in the observable features of a person - his physical as well as his psychological features - there is some intrinsic substance that always remains the same in him and accounts for his identity through time. Like Reid, he complained<sup>12</sup> that though Locke suggested this fact (in his definition of person and substance) he did not assert this (because of his insistence that the same person is not necessarily the same substance). It is presumably the supposed presence of this substance or self in a person and the absence of anything like that in an ordinary material thing that led Butler to make the distinction he made. Since such a self cannot be known in the way other persons (selves) and things are known, the standard explanation is that it is known in some special way. Butler's suggestion is that it can be known by everyone by "turning our thoughts upon ourselves".<sup>13</sup> That means that anyone who thinks that there is such a self in a person and that the identity of this makes the identity of the person is obliged to say that this self and its identity can be known, if at all, by some sort of non-sensuous intuition. Now, granted that the knowledge of the self and of its identity can be had in this special sense, it cannot be claimed that this knowledge is our criterion of making

personal identity judgments, for the self is not observable\* in the sense required for it to be a criterion. Nor can we say that the identity of the self is what we mean by saying that a later person is the same person as an earlier person: whatever temptation we may have for saying this in our own case, we cannot justifiably<sup>14</sup> say this while making identity judgments about other persons. But then it would be absurd to say that we mean this even in our own case, for this would commit us to the greater absurdity of saying that we mean different things, when we say that <sup>a</sup>later person, is the same person as an earlier person, depending on whether or not the person in question is oneself. In the face of this difficulty the substance-theorists have the only recourse of saying that the self is something simple and unanalyzable and that personal identity, which according to them is the identity of the self, is also unanalyzable. If what is said in the foregoing is right then, since Butler and Reid wanted to maintain as against Locke that selves (persons) are substances, this might be one of their reasons for claiming that personal identity is indefinable. I am not sure that this was their reason, or even one of the reasons, for making the above claim. For one thing at least, that as noted before (Ch 1 sec ii), Reid maintained that identity in general - and hence not merely personal identity - is simple and indefinable, though neither he nor Butler would want to say that the identity of ordinary material things (which surely is included in "identity in general") is due to anything permanent and substantive. (Indeed, it is because of the lack of the latter that they described the identity of those things to be imperfect.) However, an interpretation on these lines seems to be quite plausible and is not entirely ruled out.

---

\*For a fuller argument see Ch1 sec ii, see also last section pp 68-69 above.

On the basis of the presence and absence respectively of something permanent and unchanging both Butler and Reid had described the identity of persons as "real" and "perfect", and the identity of other things as "imperfect" and as "something which for convenience of speech we call identity". This is the most unambiguous and most explicit point of their thesis, and as such is more vulnerable and less defensible. For it postulates that the problem of identity - at any rate, what they called real identity - is one of finding something unchanged to account for the identity of persons and other things; but this is to misconstrue the problem at issue since, as we have seen, nothing needs to be unchanged in order that anything may be (be called) the same. However, a somewhat less vulnerable explanation can be given, and has in fact been given,<sup>15</sup> It is that our knowledge of personal identity, obtainable as it is in knowing the nature of self-identity, can be non-criterial while our knowledge of the identity of other material things has to depend on criteria. It is criterion - independence, in the described sense, that was supposed by Reid and Butler to make personal identity 'perfect identity'. And the fact that the identity of other persons, like the identity of other material things is criterion -dependent explains why the paradigm case of personal identity was thought to be self-identity, and not other people's identity. Nevertheless, the identity of other persons was not, for that matter, considered to be imperfect. One reason at least for this, presumably, is that the meaning of personal identity, being comprehensible from one's own case, could not be reasonably altered in its application to the case of other persons. Even if we may have to depend on some criteria or evidences in order to know the identity of other persons, yet that was not supposed to interfere with the meaning of "same person". As Reid says, "But still it is true that same person is perfectly

the same, and cannot be so in part, or in some degree only". Although certain observable phenomena - remembering in particular and similitude in general - were accepted as evidences or criteria for making identity judgments about other persons, scrupulous care was taken not to confuse these criteria with the meaning of personal identity. Reid, indeed, accused Locke of making just this confusion by insisting that personal identity must consist in consciousness.<sup>16</sup> This distinction between the evidences and the meaning of personal identity was no doubt a healthy sign; but the merit of this insight was overshadowed by an extreme abhorrence to criteria, which had the effect of leading the problem astray. For as we have seen (Ch 1 sec 1), the problem of personal identity is largely a problem of specifying the criteria for the identity of persons, and the theories under discussion consider the existence and use of any criteria to be irrelevant and even detrimental to the nature of real identity.

It seems that in the theories of Butler and Reid, two factors were responsible for this abhorrence to criteria. First, intuitionism. Being persuaded to think that personal identity is something intrinsic and can be known only by one's being conscious of or by "intuiting" one's identity with a past self, these philosophers reasoned that the standard knowledge of such identity must not, in principle, be mediated by criteria, and that where our knowledge of identity of something is thus mediated the identity of that thing must be declared as unreal. The second factor that might have led to this lack of concern for criteria is these philosophers' concern with the meaning of personal identity. Thus nothing is of any interest if it does not give the meaning of personal identity.

Although Butler and Reid would grant that the observable phenomena noted

above are our criteria of personal identity, yet they were of little concern. For these criteria cannot constitute<sup>17</sup> personal identity and cannot give the meaning<sup>18</sup> of it. This is true and Locke was certainly wrong in trying to define personal identity in terms of an evidence thereof. But Reid and Butler were no less wrong in so far as they implied that since the evidences, being evidences, cannot define (constitute, express the meaning of ) personal identity they must be worthless and even detrimental to the nature of identity. And I suspect that both parties were equally wrong in taking the problem as one of definition and this mistake in the one case was aggravated by the further mistake of defining personal identity in terms of what was only an evidence, albeit a ~~ex~~ criterial evidence, for it while in the other case the mistake was supplemented by the further mistake of disregarding the relevance of criteria. This obsession with meaning and the consequent abhorrence to criteria was expressed in Reid's pungent contrast between "personal identity as that which is perfect" and "the natural measure of that which is imperfect"<sup>19</sup>. But even if it is the meaning of personal identity that one is looking for, then an important way of getting at, and understanding, its meaning is by considering how the words "same person" are used. And though in most cases these words are used on the basis of the evidences or the criteria we have considered, it is these criteria, dressed up in the rather inferior terminology of "natural measures", that is not detested by our philosophers. Instead, their search for the meaning of personal identity was "inwardly" directed to bring out the result that personal identity is something ultimate and unanalyzable - with its explicit commitment to the rather dubious belief in spiritual substances.

Section (iii). The Views of Kant and William James

If anything in connection with personal identity largely engaged the attention of Kant and William James it was the notorious problem of a spiritual substance. This substance most often passed for the basis of personal unity and the bearer of personal identity through time. It will be seen that both Kant and William James agree on the denial of a permanent self or soul and also on denying that the supposition of the latter is necessary for explaining the nature of personal unity and identity.

Kant attends to this question in his chapter on the paralogisms<sup>1</sup> especially in his account of the third paralogism - where he is trying to expose the illusions of the rational psychologist with regard to the nature of what are called our selves and of our identity through time.

The rational psychologists, deriving their origin from Descartes, are known to have held an essentially non-physical view of person: person is equated by them with an immaterial, thinking substance which is also called the soul.<sup>2</sup> In Kant's picture of him, the claims of the rational psychologist amount to three main contentions: that the soul is a substance, that it is simple and that it is numerically identical throughout the different times at which the person is said to exist.<sup>3</sup> In my construal of the problems, the first two can be taken as answers to the Unity- question and the last is more directly concerned with the question of reidentification of persons. It is the existence of a substantive soul that was supposed to supply the answer to both the questions. Kant's entire effort in the paralogisms chapter was to expose the illusion that might have led to such a supposition. It is indeed an undeniable fact of self-knowledge that I ascribe to myself various experiences and thoughts

that occur not only at the same instant of time but also at different times. All these thoughts and experiences are most naturally, and without any constraints whatsoever, are addressed as mine rather than anybody else's. In this sense, they are a set of unified experiences which require "me"(or an "I") as the logical presupposition of their unity. To be aware of these experiences is to be aware of myself and, since some of these experiences may be non-contemporaneous, of my identity over time. Thus, self-identity is known, or rather expressed, in self-knowledge; and since the awareness of these experiences is a fact, self-identity is real. But, as we have seen earlier (Ch 1 sec ii) , if this is true of me because I am a person, this is also true of everyone else for the same reason. Thus it follows, pace Hume, that personal identity is real. But what makes it real, Kant would argue, is not the existence of a substantive self as the rationalists have held. Because the latter does not follow from the fact of the unity of my experiences. What can be said to follow from this fact is that the "I" in my consciousness is a logical presupposition of what I am conscious of - that it is a mere "formal condition"<sup>4</sup> of the unit of the experiences. But from this "logical meaning" of "I" nothing follows to show that I encounter a real "I" - a permanent substantive soul, which may be said to make me a single being and a being identical at different times. To jump, as the rational psychologist does, to such a conclusion is to fall prey to that fatal confusion which Strawson described as the confusion between the unity of experiences(which is a fact) and the experience of Unity (which is not a fact).<sup>5</sup> Strawson rightly pointed out that this is how Kant exposed the illusion of the rational psychologists. As a matter of fact, Kant's analysis of self-knowledge and its source, the inner sense is fundamentally empiricistic. Rather like Hume, he argued that "in inner intuition there is nothing permanent"<sup>6</sup> and that "the "I"

is merely the consciousness of my thought".<sup>7</sup> Indeed this last sentence was originally joined to the first by a "for" in Kant's passage and so Kant's intention there can be better explained by explaining the import of this. As noted earlier, Kant is willing to grant that the function of the "I" in self-knowledge is a logical function; it has got to be the logical subject of all my thoughts and experiences (Kant's word is; "representations"). In this sense, Kant calls it the "subject self"; and it is this aspect of the self which is said to be, and which has got to be, identical throughout the time during which the experiences and thoughts occur to me. Thus far, Kant is in agreement with the rationalists and is perhaps willing to stress this point to claim, against Hume, that personal identity is real and not fictitious. But unlike the rationalists, he does not accept that the self has a content of its own - much less a permanent stuff that goes to make it. By the 'subject self' Kant means simply the faculty of thought itself which is quite empty of content - "a quite simple representation".<sup>8</sup> If what is known in inner sense is just this subject, that is tantamount to saying that nothing in fact is known. The same conclusion can be derived if the self is considered in its other aspect in which Kant calls it the 'object-self'. The self in this sense he agreed, can be known in inner intuition just as the objects of sense are known in outer sense; but when it is thus known it is known not as the subject but as an object - as any other object. And whatever can be known about the nature of this object of inner sense, it cannot be anything permanent, as Hume had shown; and, as Kant also adds, "this self contains a plurality of determinations."<sup>9</sup>

Thus, as Kant saw it, the self as the subject can only be thought of and must be thought of as the formal condition of unity and identity of persons; but it cannot, as such, be known as an identical abiding substance to account for such unity and identity. Although the first part of this claim would be welcome to the rationalists and their sympathisers, the second is likely to raise some amount of indignation. It has, in fact,

been argued<sup>10</sup> (against Hume's similar conclusion) that the very assertion that I know my thoughts and experiences is inconsistent with the denial of my known myself. For, so the argument goes, to know that I am thinking or having certain experiences is to know myself as thinking or having the experiences. We have already seen (see sec 1, pp 66-67 above) that this argument is mistaken, being based on a fundamental confusion between the technical sense of the word "self" and its ordinary, non-technical sense. And if this is so, then the ordinary assertion that someone knows himself to be having certain experiences or doing certain things does not show the self in him. Thus, these ordinary assertions, which involve the use of 'self' in its non-technical sense will not affect Kant's reasons for questioning the existence and knowability of a self as the subject, and he would still be substantially right in questioning the latter. Yet he did not despair, as Hume did, as regards the identity of persons. We shall see later on in this section <sup>that</sup> this was because of Kant's merit, and Hume's failure, to have seen <sup>the</sup> importance of the fact that nothing need be changed in order for anything to be the same. This is why, the lack of knowledge, and even the impossibility of any proof of an unchanging soul-substance did not disturb Kant. He on the contrary, goes on to maintain that the identity of the self as the subject of consciousness is warranted in the very fact of this consciousness in one's own case. He gives expression to this by saying, "In my own consciousness ... identity of person is unfailingly met with."<sup>11</sup> It is not difficult to see what Kant meant by this. We have argued in earlier sections of this chapter that so far as it concerns the across-time identity of persons, "consciousness" must be understood in the sense of remembering one's past. Then what Kant can be construed as saying here is that personal identity is presupposed by the first person memory statements that one can make. This is the stronger thesis regarding the connection between memory and personal identity which

I, in earlier sections, have distinguished from the weaker thesis that one's belief in it, is expressed in the first-person memory judgments that one could make. The first thesis takes "memory" in the strong sense in which the identity of the rememberer is logically implied; the second thesis takes memory in the weak sense and does not imply the identity of the rememberer with the witness of what is remembered, yet it can be said that the relation between the relevant memory-judgment and judgment of personal identity is a logical or conceptual relation since, no matter whether or not I am mistaken in making a memory claim, I would be contradicting myself if I said that I remember having done x and that I did not do x. It is this logical connection - or at any rate the necessary conviction of one's identity which is expressed in one's first-person memory judgments - that led Kant further on to say that "we must necessarily judge that we are one and the same throughout the whole time of which we are conscious." (A364). Kant did not specify in which sense of 'memory' he wants us to understand "my own consciousness". But it will be unfair to say that he was totally unaware of the by now familiar distinction between the two sense of "memory", perhaps it will not be unreasonable to say that (being quite aware of the distinction) he has made cautious provision for mistaken memories which are now described as "weak" or mere quasi-memory.\* For he was quick to add: "We cannot, however, claim that the judgment would be valid from the/stand point of an outside observer " (A364). The observer, by applying the objective criteria of personal identity - whatever they are - can show that I am misremembering and hence that I am not the same person (i.e. the past witness) as the person I claim to be. Bennett<sup>12</sup> has rightly

---

\*I call it mere quasi-memory in order not to ignore the fact that in a sense all memories (including "strong" memory) are quasi-memories, though the reverse is not true. However, I need not say anything more about this here, though I shall return to it at a later stage.

credited Kant with this view of the matter. But he reproached him for claiming that "we must necessarily judge that ..." and (for that reason in Bennett's opinion) for "stubbornly insisting", in the face of any evidence, that we are the same person who is remembered to have done x. The truth, however, seems to be that Bennett's first accusation is true but ineffective, and his second accusation is hardly true at all. It is true that Kant said that in our own consciousness (i.e. act of remembering) 'we must necessarily judge that we are one and the same person ...' etc. and that in it our identity is "unfailingly met with". But in the light of our weaker thesis outlined above as well as in view of what Kant had said about the formal condition of the unity of our thoughts and experiences, Kant is justified (his claim is innocuous, at any rate) in saying so. For what this claim amounts to is simply that our assertion of first-person memory statements carries with it the belief in, the conviction of, our personal identity; and as far as I am concerned, this conviction remains unshaken, until at least I have been shown to be mistaken; after all the state of cognition in this case is qualitatively similar to that in the 'stronger' case. This much, to say the least, is the import of Kant's claim that "we must necessarily judge that ..." and that in our consciousness our personal identity is "unfailingly met with". But by this Kant never meant, and must not be understood to have meant, that what I thus judge is necessarily right. It is noteworthy that Kant never said anything to the effect that we must judge that necessarily.... The assertion 'we must necessarily judge that P' is certainly equivalent to 'we cannot fail to judge that P', but is not equivalent to 'we must judge that necessarily P'. From what Kant said the last does not follow - particularly if P is the proposition that I am the same person as the one who I claim to remember being. Far from implying the last Kant expressly admits that I may be mistaken and could be shown to be so by an outside observer; there is also his repeated warning that we must not

be led astray by the fact that "in my own consciousness ... identity of person is unfailingly met with". Bennett says that he can "make nothing of" this warning;<sup>13</sup> but it is not difficult to see that this warning, coupled with Kant's reference to the observer's standpoint as the standard of validity of personal identity judgment,\* give enough evidence that it was not Kant's intention "to stubbornly insist" that what we judge is necessarily right. However, if the P (above) is an assertion with regard to the identity of myself at different times, then Kant would say that we must, in our consciousness, judge that P and perhaps we should judge that necessarily P; for this, for Kant, is the formal condition of the unity of my thoughts and experiences. But Kant's warning referred to above and his similar warnings not to be misled by "the concept of personality" or by "the identity of consciousness of myself at different times" point<sup>to</sup>/his consistent contention that although these concepts (which all point to the logical subject of my experiences) are necessary for the possibility of any operation of thought and although the outside observer will admit, for that reason, "the 'I' which accompanies ... all representations at all times in my consciousness", (A362-3) yet none of these concepts implies the objective permanence of a self for the 'I' to refer to. Not only the outside observer cannot infer this, we ourselves are unable to prove this, for "we are unable to prove that this 'I', a mere thought, may not be in the same state of flux which, by means of it, are linked up with one another" (A364). It seems quite evident from this as well as from his general criticism of the paralogisms, that Kant is trying to dispel the idea of an unchanging substance as the explanatory factor of our unity and

---

\*Bennet also recognised that Kant equated "objectively valid" with "acceptable to an outside observer" (See Ibid P 96). This is yet another substantiation of my contention that the problem of personal identity is essentially a problem about other person's identity (see pp 10-11 above)

identity. As noted earlier, this puts him in a better position than many others in the sense that he is free from the dogma that "sameness" and "change" are incompatible. The following significant passage in his Anthropology gives a clearer proof of this:

"The question may be asked whether, in view of the variety of changes of mental states ... a man be conscious of these changes and still say that he remains the same man (has the same soul). The question is absurd, since consciousness of such changes is only possible on the supposition that he considers himself in his different states as one and the same subject".<sup>14</sup>

Thus according to Kant, one can consider himself the same person (which, needless to say, Kant used as equivalent to soul or subject) despite his knowledge of the changes (which is all that he can observe both in his outer sense and in his inner sense), and yet the subject or the soul is nothing permanent and abiding. Moreover, Kant is even prepared to entertain the hypothesis that what is called our soul is not something simple and unitary (Critique of the 2nd paralogism) but a plurality of souls, each soul transmitting its conscious states to its successor soul so that each soul could be conscious of all that had happened to its ancestor souls, without being identical (numerically) with these others. This speculation comes out in a remarkable passage in the Critique:

"If ... we postulate substances such that the one communicates to the other representations together with the consciousness of them, we can conceive a whole series of substances of which the first transmits its state together with its consciousness to the second, the second its own states with that of the preceeding substance to the third ... The last substance would then be conscious of all the states of the previously changed substances as being its own states ... And yet it would not have been one and the same person in all those states." (A363-4 note)

This hypothesis was to be of considerable significance to later theorists of personal identity; and we will presently see how this insight influenced James' theory of personal identity. But for Kant, however, this was just an empty speculation entertained, only for arguments sake, to discourage the rationalists speculation of a soul-substance. And as regards the

nature of personal identity Kant did not give any answer that is either clear or categorical. I can think of the following reasons for this. He equated persons with souls. These souls, he argued, can only be known, if at all, in inner sense; but the inner sense, as well as the outer sense, can only give us the knowledge of things not as they are but as they appear to us. (This is the principle of Kant's phenomenalism). So, in whatever manner the soul is known in inner sense it cannot be known as it is in itself. (This is the inevitable agnostic off-shoot of his phenomenalism). This latter, incidentally, is Kant's "real" self as opposed to the "phenomenal" self, which can be known in inner sense. But the "real" self can only be an object of thought, not of knowledge.<sup>15</sup> It followed therefore that the real nature of personal identity (being, for Kant, the same as the identity of the soul) cannot be known.

Nevertheless, the concept of "personality" must be retained and is necessary "for practical employment and is sufficient for such use".(A366) But the personality that is thus retained is still conceived in the "internal" model and confined to the inner sense as a knowable object. This model of thinking did not allow its author to consider the important fact that persons, at least in the "practical employment" of the concept, are importantly - if not entirely - objects of outer sense.<sup>16</sup> It also leads Kant to ignore the equally important fact that in our practical employment of the concept, we have empirically applicable criteria of personal identity, which though not the same as bodily identity, have "an essential reference to the human body".<sup>17</sup> But by giving very little consideration - almost none - to the bodily criterion, even to the very fact that persons have bodies, Kant fares no better than his rationalist adversaries. His polemics are no doubt instructive, but what he achieved was hardly ever so.

Like Kant's, William James' theory of personal identity is marked by a conspicuously critical attitude towards the idea of a substantive soul; and both men seem to have derived the force of this attitude from Hume's argument against the idea of the soul. But whereas Kant emphatically believed in the soul as the logical condition of unity and was hesitant about its nature, James straightforwardly rejected the idea of the soul as superfluous for any scientific and useful purpose and also on the ground that there is no reason to believe in it, but was more categorical and committed as to the nature of what is called the soul: as a matter of fact he had no hesitance to hold a view treating the soul as only a "bundle of Humean perceptions". One can see clear reflections of the basic philosophical assumptions of these philosophers in their respective treatment of the problem of personal unity and identity: in the one case, as we saw, agnosticism, in the other pragmatism.

From a rather more scientific point of view, James says, "One great use of the soul has always been to account for, and at the same time guarantee, the closed individuality of each personal consciousness. The thoughts of our soul must unite into one self, it was supposed, and must be eternally insulated from these of every other soul". (349)<sup>18</sup> But James does not think this to be of any advantage. His immediate reason of course, seems to be the possibility, which he seems willing to allow, that "in some individuals, at least, thoughts may split away from the others and form separate selves." (349-50), so that, presumably, the individuation of personal consciousness cannot be guaranteed by the sameness of soul or self. But without needing to entertain this rather dubious hypothesis to explain such abnormal phenomena (which perhaps, could be better explained otherwise), it is not difficult to see James' point as well as of his general "anti-substantialist" attitude. For the "substantialist view", which springs naturally from the supposition

discussed above, fails to explain anything. We have noted Kent's contention that the only sense in which a soul-substance can be justified - if at all - is that it can at best be taken as a "formal condition" of the unity of our thoughts and experiences. The constant talk of "I" (me) as the subject of my experiences and the reference to all these experiences as my experiences makes my identity (or the continuous presence of myself, as it is usually expressed) a necessary presupposition of such talks. But this is only a verbal point\* from which nothing can be deduced about the nature and content of this "I" - much less that it is a permanent soul-substance. This would be as absurd as saying that any word that can be used as a grammatical subject in a sentence must have a reference in order for that sentence to be meaningful. The absurdity of such suppositions is fairly well-known; equally well-known is the fact that the substantialists have been arguing on some such basis. Moreover, after having thus erroneously deduced this unknown something, and giving it the name of 'soul' the substantialists proceed to explain the nature of our mental phenomena by its means. But the nature of the soul itself being unknown and even unknowable, it fails to meet the requirement which its authors wanted it to meet. Any account of its nature is bound to be given - if at all - in terms of, or at least on the model of, the very same subjective phenomena which it was designed to explain. As Ayer has succinctly put it, the soul can be represented - if at all - "as a transcendental stream of consciousness, duplicating the one we know".<sup>19</sup> James himself has further, more genuine explanations to show that it is superfluous and is not required "for expressing the actual subjective phenomena of consciousness as they appear", (344) which it purports to do.

---

\*Kant of course would not say that it is a verbal point, but I do not see how consistently he could deny this.

We must now turn to these explanations which rather show a novel feature in his approach.

James tried to present his case against the "substantialists" by his famous theory of appropriation, which he pleads and elaborates in the section on "The Pure Ego" in the 10th chapter of his Principles of Psychology. One very important argument for the supposition of a soul-substance has been that experiences must have a owner. Although, the idea of an unowned experience is not straightforwardly self-contradictory,<sup>20</sup> it was thought to be at least counter-intuitive since commonsense construed experiences as being the experiences of someone; this latter fact was supplemented by the almost universal fact that experiences are referred to as mine or yours etc. Hence it was supposed that if there is an experience there must be an owner and the soul was invented to play the owner and to own the many different experiences that are said to belong to a person simultaneously and successively. The Humean tradition, in its eagerness to deny the idea of soul, had denied the idea of ownership altogether and had let loose the experiences - with 'no real bond' to bind them together. The effect was the unreality of personal unity and identity. James, on his part, though in essential agreement with the Humean tradition, tried to repair the "loss" by suggesting that the ownership, instead of having to depend on the soul, can be said to belong to the experiences themselves. This he explains by saying that what we call our "self" might be very well construed as a series of momentary selves\* each one of which transmits all

---

\*James own word was "thought" with a capital "T". To avoid the monotony - if not the oddity - of talking of "thought" appropriating thought and being ~~appropriated~~ by "thought", I have expressed James' view in terms of 'self'. I do not think that this practice would mean any alteration in, or, offence to, the purport of James' theory, and I hope that my 'self' would perfectly fit in. Moreover, James made it clear that if he may have to use this word, "self" should be taken in its popular sense with no substantialist implication whatever. (350). Although my use of 'self' here on James' behalf, cannot be as popular a sense, it is at best a semi-technical term acceptable to James and designed only to do the same work as his "thought".

its consciousness to its successor self which 'adopts' this consciousness as its own and, in its turn, transmits its own consciousness together with what it had adopted from its earlier self, and so on. The result is that the last self is conscious not only of what happened to it alone but also of all that had happened to its predecessor selves - both alike being felt as its own. This process of owning or 'adopting' whatever happened to the previous selves is what James describes as the process or act of "appropriation". The idea is that every self in the series is born an owner and dies owned, "transmitting whatever it realised as self to its own later proprietor" (330). This insight is one of several<sup>21</sup> of James' indebtedness to Kant. He writes:

"Kant says, it is as if elastic balls were to have not only motion but knowledge of it, and a first ball were to transmit both its motion and its consciousness to a second, which took both up into its consciousness and passed them to a third, until the last ball held all that the other balls had held, and realised it as its own. It is this trick which the nascent thought has of immediately taking up the expiring thought and 'adopting' it, which is the foundation of the appropriation of most of the remoter constituents of the self. Who owns the last self owns the self before the last, for what possesses the possessor possesses the possessed". (339)

Now if this theory of 'appropriation' or 'adoption' of experiences by experiences is correct, then the doctrine of ownership is retained and explained without the supposition of unchanging soul-substance and the reality of personal identity is restored. For a person at any moment can be judged to be identical with the witness of whatever he is then conscious of since he, being the "proprietor" of those experiences, is conscious of them as his own. (This psychological analysis appeals at least more plausibly to our logical intuitions than the metaphysical theory of a permanent soul). And what is called the same person need not be one identical substance. James' theory of personal identity, thus, seems to be quite similar to Hume's with the essential modification that personal identity is claimed to be real and not "fictitious". Hume's despair was

due to the fact that "all our distinct perceptions are distinct existences and the mind never perceives any real connection among them".<sup>22</sup> James' improvement upon this is that although it is true that our perceptions are distinct in the sense that they are logically independent of each other, it is not true that there cannot be any real connection among them. As Ayer expressed the point "That they (our perceptions PKM) are separate does not entail that they are disunited."<sup>23</sup> What can be a more real connection, James would probable ask, than the fact that our perceptions appropriate and are appropriated by one another in the sense outlined? This factual connection is also expressed by saying that "within each personal consciousness thought is sensibly continuous" and that changes from one moment to another in the quality of consciousness is never absolutely "abrupt."

James' theory thus has the merit of making our perceptions do the work for which others had to invent a separate agency. And in a sense his whole theory seems to depend upon this hypothesis. Unfortunately this is not as sound and legitimate as one would wish it to be. The greatest stumbling block is the concept of 'appropriation' itself. Even James himself conceded that it is "the only point that is obscure" (340) in his theory. James has two ways of explaining how this act of appropriation may be possible. One way is by saying that an identifying thought (a present self) appropriates whatever experience it feels "sensibly continuous" with itself; the second way is by saying that it appropriates whatever experiences are marked in its recollection with "warmth and intimacy" (331). It may be pointed out, in parenthesis, that in trying to define personal identity, or at any rate the unity of experiences, in terms of the act of appropriation James is committed to a 'memory-theory' of personal identity which is why there is little allusion to and less emphasis on the bodily criterion in his theory. Now, besides being a matter of feeling 'sensible continuity' can hardly be a necessary condition

of personal identity. For if it is so taken then in the case of an abrupt change of memory etc, like that in amnesia for instance(compare the theory of Locke, section(i) above), the person after the change has to be different from the person before the change. But this is surely counter-intuitive and to come to accept - if we must - such a conclusion would at least involve taking a decision. Further, that "sensible continuity" cannot be taken as a sufficient condition either is evident from the fact that we often ignore discontinuity, or even total loss, of memory and ascribe identity to persons on the basis of bodily continuity. James also admits this fact in a crucial passage where he says:

- "... even were Thought entirely unconscious of itself in the act of thinking, those 'warm' parts of its present object(which he equates with the body and what he called the central adjustments,(PKM) would be a firm basis" of personal identity(341).

Thus, in view of the above, the relation of "sensible continuity" can supply only a very weak explanation of 'appropriation: The feeling of "warmth and intimacy" fares no better either. It reminds us of the 'feeling of familiarity' and 'vivacity' which Russell(at one time) and Hume respectively held to be constituting the distinctive feature of memory. But all these, being mere matters of subjective feeling, can never be the necessary accompaniments of personal identity - not even of memory. Sometimes some objects of imagination may seem to be more 'familiar' and may feel more "warm and intimate", shall we have to say that we are the same persons we imagine ourselves to be ? James could provide no protection for his theory of appropriation against such eventualities. Moreover "warmth and intimacy" being inevitably subjective, are bound to be matters of degree; and depending, as he does on these factors, James can never escape falling into the odd way of treating personal identity as a matter only of degree.\*

---

\*The oddity of this way of talking about personal identity has been indicated in Ch 1 sec 1. More will be said on this in the next Chapter.

It follows therefore that the concept of appropriation remains an admittedly obscure concept and James' theory of personal identity, depending heavily on this concept, remains far from satisfactory. And his primary emphasis on the appropriation-hypothesis committed him solely to the inner criteria, despite his apparent willingness (as he implied in the passage last quoted above) to say that we all identify ourselves with our bodies as persistents.

Chapter 3

THE CONTEMPORARY APPROACH

Prologue:            A Problem of Criteria

Our analysis of the traditional theories has enabled us to reveal an important feature with them; namely, a dominant tendency to take the problem at issue~~as~~ as one of definition. This tendency, resulted, in most cases, in defining personal identity in terms of what only was a criterion thereof, and in a few cases, where no definition was to be found, the result was intuitionism. This tendency and the consequent difficulties do not as much vitiate recent writings on the subject and there is a clearer tendency in the latter to keep the "meaning question" and the "criteria question" apart: \* this will be brought out in the subsequent sections of the present chapter as we proceed.

Now, one more thing before we pass on to the next chapter. The difficulties issuing from the described tendency have been demonstrated in the case of the psychological criteria of personal identity. Any attempt to define personal identity in terms of the other criterion, that of bodily continuity, will be equally open to similar difficulties. A straightforward way of demonstrating this difficulty is that if personal identity is defined in terms of bodily continuity, then the continuity of the same body would logically entail the identity of the person. But this is prima facie absurd, since the same body does continue (after what we call the death of a person), at least for some time,

---

\* I cannot pretend, though, that this is true of all contemporary writings on the subject. For there is at least one recent theory ( See sec iii below) which presents a glaring exception.

without there being any person to be the same person as the earlier person who was alive in that body. If, alternatively, it is argued that the person still continues to be there in a disembodied state (which is not logically impossible), then at least bodily continuity cannot be a defining characteristic - much less the definition - of the identity of persons. Further the logical possibility of disembodied existence of persons will count against a logical definition of 'same person' in terms of bodily continuity. Nor can the view that 'same person' can be defined in terms of the continuity of a living human body fare any better. For as is evident from a recent paper by Brian J. Smart,<sup>1</sup> although the same living human body continues from the stage of foetus to the stage of adulthood, say, yet there was no person in the former stage to be judged identical with the person of the latter stage. 'The same living human body' can serve to define personal identity if and only if (1)  $P_1$  and  $P_2$  are both persons and (2) both  $P_1$  and  $P_2$  have the same living body. But as Smart's example shows, although condition (2) is satisfied, condition (1) is not satisfied. This argument may not knock down the status of bodily continuity as a necessary condition of personal identity, but it certainly would not allow the latter to be defined by bodily continuity (since a definition of x is the necessary and sufficient condition of x). To this can be added the further argument, advanced notably by Hick,<sup>2</sup> that a person may (it is logically possible or conceivable) instantaneously change bodies without becoming a different person. This possible phenomenon points to the possibility that there can be personal identity without there being a continuous living human body (see note 2 now).

Thus it follows that any attempt to define personal identity in terms of the one or the other criterion thereof is bound to fall short of a logical definition. (That personal identity cannot be defined in terms of

both the criteria together has been shown earlier in ch 1 sec i ). I submit that those difficulties will not arise if similarity of memory (speaking very generally) and/or bodily continuity are taken as criteria of personal identity. This will be a fairly innocuous but useful approach to the problem; for as we shall see in what follows, a criterion does not entail the existence or occurrence of what it is a criterion. Much of the difficulties noted above is due to this entailment that holds between the definiendum and the definiens. A criterion, by contrast, will be said to be something which will not entail but which, nonetheless, will justify the existence or occurrence of the phenomenon in question. Thus construed the role of memory and bodily continuity will be seen in their proper perspective and the problem at stake will not be led astray. What, then, is a criterion and what sort of phenomena will serve as criteria for something or some state of affairs ? To this question we shall turn now.

Section(i): Criterion\*

Earlier, I had described a criterion as a state of affairs the existence of which necessarily is evidence for the truth of the judgment(s) of which it is a criterion (see p 25 above). Although I believe that this

---

\*The term "criterion" is used very widely, and I shall so use it, so that a phenomenon or a state of affairs may be a criterion for the existence of another phenomenon or state of affairs. It is also used in such a way that a statement or a judgment may be said to be a criterion for another. In the latter way of speaking, what is meant is that the state of affairs described by the first statement (or judgment) is a criterion for the state of affairs described by the other statement (or judgment). Properly speaking, nothing can be said to be a criterion for a statement (or a judgment) but only for the truth of a statement (or judgment).

f gives a fair account of what a criterion is, a lot more needs to be said to bring out the purport of this description. Of course "criterion" is a technical term and can be defined as one wants it to be. But the way I want to use the term will not be very different from how most philosophers have wanted to use it. I therefore choose to begin by indicating the dominant features of a criterion, and then specify the sense in which I want to use it. What this dominant feature amounts to can be best brought out by considering some examples of what standardly would be called criteriological relations(hereafter, c-relations). Such relations are said to hold between pairs of propositions like the following:

Some-one is displaying(what we call) pain-behaviour and he is in pain.

Someone is muttering in sleep, his face changing expressions as he does so and he is dreaming.

Someone gives correct answers to most questions in a quiz and he is intelligent

Each of the former propositions in this illustration is said to be a criterion for the truth of the corresponding latter proposition. This means that under all normal circumstances we can say that in each case the latter proposition is true if the former is true. The qualification "under all normal circumstances" underlines the fact that a c-relation is not one of entailment, for it is fairly conceivable—even factually possible — that in any particular case, though the former proposition is true the latter is nevertheless false. But although the relation is thus not strictly logical, we are justified in saying that any of the latter proposition is true if the (corresponding)

former is true, and we have no justification for saying that the latter is true if the former is not true - though nevertheless the latter may be true. In other words, a criterion for a given thing's being so is something by which one is justified in saying that the thing is so and in whose absence (coupled with the absence of any other alternative criteria) one has no justification whatsoever for saying that the thing is (or even, is not) so. However, the fact that what is called a criterion may sometimes fail to give the right result need be no reason for saying that it is not a criterion under all conditions or that something's being a criterion ~~as~~ is relative to some conditions only. For one thing, if it fails to give the right result the question is: what fails to give the right result? It is the criterion that fails. We do not say of any phenomenon but only of the phenomenon of what is called pain-behaviour that it gives or fails to give the right result, namely, that the subject is in pain. Therefore, if p is a criterion of q it is a criterion, simpliciter, of q; and the "abnormal" conditions (under which p naturally gives the wrong result) are called "abnormal" because under these conditions p (the criterion) fails to give the right result. We say of a case of someone's pretending to be in pain that it is "abnormal" because despite the phenomenon of pain-behaviour, he is not in pain. The fact, therefore, is that if something is a criterion for some-thing else then it is always so - though its giving the right result or wrong result depends on whether the circumstances are normal or abnormal. Moreover, the fact that we can make mistakes sometimes while working on the basis of a criterion cannot discredit the criterion or make it cease to be the criterion.

This point can be brought out by taking an analogous illustration from Pollock:<sup>3</sup> the fact that someone applies a concept (e.g. "bird" ) to a wrong object (e.g. to a duck hunter's decoy from a distance) cannot show that he has not learned the meaning of the concept. That he has is evident from the very fact that he did apply it to something very much like a bird. In other words, his application of the concept was justified (hence his learning of its meaning accomplished), though, due to extraneous conditions, false or mistaken. To learn the meaning of a concept is hence, " to learn how to ascribe it justifiably to things"<sup>3</sup> (my emphasis). Similarly, a criterion is a "justification condition"<sup>4</sup> and the test of something's being a criterion of something else is whether we can say justifiably (even if falsely, sometimes), on its basis, that the latter is true.

It is important for our purpose to note that a "justification condition" is quite different from a defining condition and that a c-relation is not a relation of entailment. The fact would need no mention if it were not a dominant tendency among some to think otherwise, and to treat the c-relation very much like one of entailment. Roger Albritton, in giving one interpretation of Wittgenstein's use of "criterion" said:

A criterion for a given thing's being so is something that can show that thing to be so and show by its absence that the thing is not so; it is something by which one may be justified in saying that the thing is so and by whose absence one is justified in saying that the thing is not so<sup>5</sup>. (The first two emphases mine)

In bare outline this passage makes it appear as if the c-relation is no different from the relation of entailment. But that this latter is mistaken will be evident by considering the passage in its finer detail. The passage may be understood to express two theses: an amplified strong

thesis,\* that a criterion for a given thing's being so can conclusively prove (or is conclusive evidence) that the thing is so and that its absence can conclusively show (or is conclusive evidence) that the thing is not so; and an expressed weaker thesis that a criterion is a justification (which, by no means, is conclusive proof or evidence) for saying that the thing is so and for saying, in its absence, that the thing is not so. Taken in the first sense, the passage is entirely mistaken (even as an account of Wittgenstein's view - as we shall shortly see), since, as we saw, not only can we think of a criterion yielding wrong results, we can also think of something's being so without the relevant criterion being instantiated; - indeed these are factually possible. Nor is it what Wittgenstein would want to say. For clearly a number of evidences listed as criteria in the Philosophical Investigations (hereafter PI) are not conclusive evidences (e.g. PI pt I sec 377, also of Pt II p 222 and pt I sec 56). And as Malcolm pointed out, Wittgenstein would "clearly" deny that a criterion is ever a conclusive evidence.<sup>6</sup> I shall submit further that Albritton's passage, taken in the second sense as indicated above, though expressing what Wittgenstein might have said is at least partially wrong. For although we are certainly justified in saying on the basis of the criterion that something is so, we have no justification, in its absence alone, for saying that the thing is not so. Are we justified in saying that Smith is not in pain simply on the basis of the fact that he is not engaged in pain behaviour? Certainly not. As we have said, the truth simply <sup>is</sup> that in the absence of pain-behaviour we are not justified (assuming, of course, that there are no alternative criteria which we know to have been satisfied and which hence point to the contrary) in saying that he is in pain although, as noted before, he may be in pain nonetheless,

---

\*For there is a sense in which "shows" means "proves" or "conclusively establishes" in the logical sense of the expressions.

but NOT that we are justified in saying (in the absence of pain behaviour alone) that he is not in pain, for he may as well be.

However, these shortcomings in Albritton's interpretation are not so much of his own makings as they are due to Wittgenstein's unclear expressions of this disturbing concept (though he (Albritton) must still be blamed for not giving any essential guidelines to carve out a definite notion of "criterion" from the Wittgensteinian obscurity). Wittgenstein's use of "criterion" is found to be oscillating between definition and inductive evidence or what he distinctively called symptoms. The Blue and Brown Books (hereafter BB) seems to make criterion almost indistinguishable from "definition". This is evident from his claim that " ... to say 'A man has angina if this bacillus is found in him' is a tautology or .. a loose way of stating the definition of angina",<sup>7</sup> and also from the fact that "criterion" is often seen to have been equated with what he called "defining criterion" (See esp. BB p 25 paras 1 and 2). Now if this version is taken seriously, a criterion would have to be both necessary and sufficient condition for what it is a criterion of and the c-relation would have to be a strict logical relation of entailment. But our analysis of the notion, and particularly the above criticism of Albritton's interpretation, has shown that this is not the case - and this is not true of what most philosophers would want to call c-relations. Incidentally, I will not rule it out that something could be called a "defining criterion". As far as I can make of this, a state of affairs could be called a 'defining criterion' for another if and only if it can be the criterion - in the sense of one and only one piece of necessary evidence - for that other. But whereas this could be a sense of the term 'criterion' it would be a very strong sense indeed: such a criterion would, in effect, be indistinguishable from definition. For in being the only piece of necessary evidence, this criterion has to logically guarantee the truth of what it is the defining

criterion of; and in this sense it will entail the latter - which is what a definition does. The curious fact is that Wittgenstein cannot be said to be committed to this. We have just noted his "negative answer" to such a proposition (see note 6); further, he sometimes gives way to such ways of talking as might draw criteria nearer to symptoms rather than to definitions.<sup>8</sup> (Though, irritatingly, he cannot be said to have been committed to this either.\*)

However, things are not as desperate as they look like. For there is something which Wittgenstein was committed to and which rightly contributes to a reasonably clear and useful account of criterion and this is the view, recently brought out by Shoemaker, that a criterion is a non-inductive evidence and that the c-relation is not an empirical relation. This explains why he was so anxious to liken criterion to definition and to keep it distinct from symptom. One of the unchanged part of his thesis consists in the following way of distinguishing a criterion from a symptom, namely that the evidential value of a symptom is something taught to us by experience (see BB p 25) whereas the evidential value of a criterion is something "founded on a definition" (PI sec 354, also cf BB p 25). But if a criterion is thus not to be taught to us by experience and yet not to be equated with definition, it must enjoy a somewhat intermediary status between definition and symptom. Unlike a definition, a criterion is not to be taken as entailing the existence or occurrence of what it is a criterion of, but it must not, for that reason, be taken to be empirically (so contingently) connected to the latter - just as a symptom is related to what it is a symptom of. This unchanged, and correct, intention of Wittgenstein was not entirely - though considerably - eclipsed by his

---

\*To be evident from the following paragraph.

various loose and obscure ways of talking about 'criterion'. In recent writings, Shoemaker has very succinctly expressed it by saying that a criterion is a "direct and non-inductive evidence"<sup>9</sup>. Thus, a criterion is a special sort of evidence which, unlike other evidences, viz, symptoms, is not inductively known to be an evidence. For it is not "what we have (experimentally) found to be evidence", but "what we have ... learned to call evidence"<sup>10</sup>. Criteria earn this "privileged intermediary status" from the essential role they play "in the way certain concepts are formed and in the way certain words are learned"<sup>11</sup>. This point is very often stressed by Wittgenstein. For example, if pain behaviour were not connected with one's having pain in the usual way, i.e. if people did not display pain-behaviour when they are in pain, no one could have taught the use of the word "pain" to anyone else and the word would not have any meaning or, at best, it would have had a different meaning (cf PI sec 385). For if the usual link between pain and what we call pain behaviour were otherwise, either people did not pain-behave at all or they "pain-behaved" under different situations (normally always\* when they felt ticklish, for example). But if people did not pain-behave at all, there would be no way in which anyone could tell or teach anyone else what pain is (i.e. what does the word "pain" signify). We are taught the meaning of the word "pain" by being shown people who are in pain (or so at least they seem) and we know

---

\*I rule out the possibility that one might display the same behaviour (viz, what we call "pain behaviour") each time under a different situation. For I think that this would not only be abnormal but also that, there being no particular phenomenon to go with that behaviour, even he himself would find no word to signify or qualify that behaviour; and any word ('pain' or something else) we and he may choose to qualify the behaviour by is bound to be without a fixed meaning. It is therefore a matter not of fact, but of logic that people normally behave in similar ways in similar situations. For this is necessary in order for the crucial word to have a meaning.

that people are in pain from what they do and say. If nobody ever behaved as if he is in pain or said that he is, nobody could have known what pain is and even there would be no word called "pain". And if, alternatively, people usually behaved as if they are in pain (i.e. like how they now do behave when they are in pain) when in fact they felt some other sensation then "pain" would not mean what it does and would mean that other sensation. It is in this sense that a criterion is necessarily tied up to meaning and is said to be "founded on a definition", and this makes it what we called a necessary evidence. But this is not to say that a criterion (or even all criteria for a given thing) is the meaning or is the definition (cf PI sec 299). We say of a man that he is in pain when he groans or moans, carefully nurses a part of his body and implores for anaesthesia and so on; but that is not what "being in pain" means, we say of a man that he is dreaming if he mutters something in his sleep and so on, but that is not what "dreaming" means (see Malcolm, Dreaming<sup>12</sup> p 60). A criterion shows what pain or dreaming is, but it does not mean what they mean (and we have already indicated how a criterion shows it). Nor, for the reasons stated above, would we be right to say that the criterion is empirically (so contingently) related to what it is the criterion of.\* And as Shoemaker has very rightly said, "If so-and-so's being the case is a criterion for the truth of a judgment of  $\phi$ -identity the assertion that it is evidence in favour of the truth of the judgment is necessarily rather than contingently

---

\*Furthermore, an empirical relation in this case would have to be at least like what holds between a symptom and what it is the symptom of. Wittgenstein calls symptom "a phenomenon of which experience has taught us that it coincided ... with the phenomenon which is our defining criterion" (BB p25). If this is so, then the fact of there being a symptom would imply that there must be something, which <sup>is</sup> a criterion, for it to coincide with. But if a criterion too has to be thus empirically related to what it is the criterion of, then it would, like a symptom, require another "criterion" to coincide with, and this, for the very same reason, would require yet another "criterion" to coincide with, and so on infinitely. Consequently, nothing would ever be a criterion; equally, nothing could ever be a symptom either.

(empirically) true. We know that it is evidence, not by having observed correlations and discovered empirical generalisations, but by understanding the concept of a  $\phi$  and the meaning of the statements about the identity of  $\phi$ 's".<sup>13</sup> Thus, being determined by our understanding the meaning of the concept, a criterion is direct, and being knowable by no observed correlations it is non-inductive evidence.

We are now in a position to see at least two senses in which a phenomenon p can be said to be a criterion of another phenomenon q.

- (1) That p is logically necessary and sufficient condition for q.  
e.g. when p could be what we have seen to be the "defining criterion" q.\*

and (2) That p is a good reason or justification for q, but is not inductively established to be so.  
e.g. the general but violable 'justification conditions', as we have called them.

Although both senses (or sorts?) of criteria above may satisfy Shoemaker's requirement (which we subscribe to) of being 'direct and non-inductive evidence' yet (1) being logically necessary and sufficient condition, is non-distinct from definition and has to involve entailment. It is for this reason we pointed out that this would be too strong a sense for the term to be used at all. We further argued that a criterion is not a conclusive evidence, but a defining criterion would be an evidence that is conclusive. We must therefore fall back upon the second sense for a standardly acceptable, usable and useful account of "criterion". In this sense a criterion does not and need not logically entail the existence (or truth) of what it is a criterion of. We have seen that much of the difficulties involved in any attempt to define personal identity is due to

---

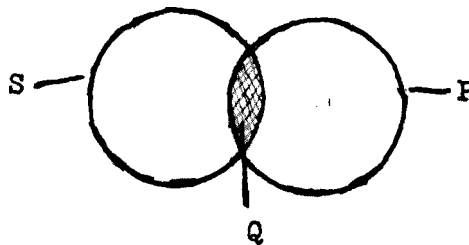
\*There is, to be sure, a difference between definition (with which I have in effect equated "defining criterion") and what is/are necessary and sufficient condition. But the point is that the latter, like a corollary from a definition, would entail; and it is in this last respect that I claim it to be indistinguishable from definition.

this entailment that is inseparable from any theory of definition (or of meaning). In contending, therefore, that our problem is not one of definition or meaning of personal identity but one of criteria thereof we envisaged this second sense of the notion of criteria: the first sense would be doubly disadvantageous for our purpose since it would bring back all the difficulties that beset a theory of definition of personal identity with.

There is just one more thing which I would add regarding my use of the concept. I have wanted to say that a criterion is simply a justificatory evidence and that it does not and need not entail the truth of what it is the criterion of. In this respect it is to be contrasted not only with a definition but also with a necessary condition. For although a necessary condition usually does not and need not entail the truth of whatever it is a necessary condition of, yet the conjunction of necessary conditions do (and even one single necessary condition - if it is the one and only one such condition may) entail. Our contention, on the contrary, is that even if all the criteria for a thing's being so are satisfied it will not entail that the thing is so. (It is, for example, perfectly coherent to suppose that a certain being looks like a man, <sup>acts like a man, in all observable respects</sup> is like a man, but is not a human being). Moreover, there is a general asymmetry between the two concepts. In the sense of 'criterion' I have outlined, pain behaviour (understood non-question beggingly) is a criterion of someone's being in pain, but as we have seen, it is not a necessary condition of the latter. On the other hand being a material object is a necessary condition of something's being a book, but it will not be a criterion for the latter. Nevertheless it does not follow that the two concepts are logically incompatible. For I think that a (not any) necessary condition of something being so can be a criterion for that thing being so as long as it is a justificatory evidence for that thing being so, i.e. as long as one is justified, on its basis, in saying that the thing is so. For example, it is a necessary condition

for someone's understanding the meaning of a word that, more often than not, he is able to use it in the appropriate context. But this very fact is also our justification for saying that the man has understood the meaning of the given word. The important thing, however, is that a fact's being a criterion for something does not (logically) depend on its being a necessary condition for that thing. I can be justified, on the basis of x, in saying that y is true, and I may, perhaps, stumble on the fact that x happens to be a necessary condition for the truth of y; but I don't have to be aware of this latter fact. On the contrary, and this is the crucial point, x may be found not to be a necessary condition of y, and yet it may continue to be our justification for saying that y is true. Later on, we will try to make plausible the idea that bodily continuity( in a sense to be qualified then), which is a fair ground for saying that a later person is the same person as an earlier person, is also a necessary condition for personal identity. But this fact will in no way restrict our use of the term "criterion" to necessary conditions only. For we will also be using the similarity of memory and/or character etc as another criterion of personal identity, since it is also equally a fair ground for saying that a later person is the same person as an earlier person; and it will also be seen that memory continuity is not a necessary condition of personal identity.

The following diagram will explain our account of criterion as related to necessary condition.



S = Justificatory evidences or criterion in our sense.

P = Necessary conditions.

Q, the shaded area, represents those criteria which happen to be necessary conditions as well. If the argument that bodily continuity is a necessary condition of personal identity would be right, the criterion of bodily continuity will fall under this category of criteria.

However, our general requirement is still that a criterion need not be a necessary condition and must not entail. Criteria like bodily continuity will be best construed not as exceptions to this requirement but rather as a sort of additions to it: it will only indicate a broader view of criteria which we wish to embrace. Even if I grant the possibility that in some cases a criterion may thus coincide with a necessary condition I contend that this is no reason to confuse, or try to identify, the two. This confusion would be symptomatic of the more general confusion of criteria and definition. For, like the latter, this confusion would mislead us into thinking that in the absence of a criterion one would be justified - ~~may~~ nay, rather logically bound - to deny the truth of what it is the criterion of. And this on my account of criteria, is a definite mistake. (In the very few cases of the described coincidence, some criteria may be "privileged" but it is important to see that not all criteria are and that not any need be ).

Section (ii) Bernard Williams: Bodily continuity

It has been seen that one enters into difficulties if one tries to define personal identity or wants to know what its meaning is. So the proper question to be asked, it was argued, is: how do we know, or what evidences can we have, that a later person is the same person as an earlier person? And our deliberations in the last section has shown that the evidence(s) sought is a special kind of evidence which, under normal circumstances, will necessarily show that the later person and the earlier person are one and the same person. Such evidences, which we have identified with what technically are called criteria, are mainly (though not necessarily only) two in the present case: bodily continuity and the continuity (strictly, similarity) of memory-claims, character, personality etc.\* Usually, two non-contemporaneous persons are judged to be identical on the basis of bodily continuity and/or of memory continuity. Considerations of the latter has occupied the prime place in the traditional theories of personal identity, so much so that the bodily considerations have got little or no place at all in these theories. The standard explanation given of this is Cartesianism. But even for those who do not rejoice over any form of Cartesianism or near-Cartesianism, psychological considerations like those of memory and personal characteristics, still figure primarily since persons are said to have minds (Cartesians are not certainly wrong in this, however wrong they may be in what they said about this) and are capable of what is called higher-order thinking and intelligence.<sup>1</sup> It is apparently for this reason, if not for the monotony of a non-physical or an essentially mental picture of persons, that the memory-continuity criterion still retains its importance in the theories of personal identity. And that is why memory-continuity is advocated to

---

\*The latter is usually referred to as the criterion of memory continuity and refers to similarity of memory-claims which may or may not be accompanied by that of character and/or personality etc.

be a strong and important criterion of personal identity. Many in fact do advocate that it is not only the primary criterion but also the sole criterion of personal identity. To such theories, bodily continuity is at best a secondary criterion.

A very fundamental limitation of such theories, which concerns the principle of identity itself, has been exposed in recent times by Bernard Williams. In a very influential article called "Personal Identity and Individuation",<sup>2</sup> Williams argued that bodily continuity is a necessary condition of personal identity and more particularly, that continuity (similarity) of memory could not be a sufficient condition of it.\* This he tried to show by pointing out the absurdity that follows if similarity of memory and character etc were taken to be sufficient condition of personal identity. He asked us to imagine that a certain person, say Charles, undergoes a sudden change and acquires characteristics which are exactly like those of a person, say Guy Fawkes, known to have lived in the distant past. Charles is to be further supposed to be making sincere memory claims which, when checked by records, entirely fit the life of Guy Fawkes. This may tempt one to say that these conditions are sufficient to identify Charles with Fawkes. But, Williams argued, they are not. For if it is (logically) possible for one person, Charles, to undergo the described change, it is also (logically) possible for another person, say Robert, to undergo simultaneously an exactly similar change. And if the conditions were sufficient to say that Charles is the same person as Fawkes they should as well be sufficient to say that Robert too is the same person

---

\*Williams has argued that memory cannot be a sufficient condition of personal identity and that therefore bodily identity is always a necessary condition. The "therefore" does not seem to follow. But it follows from that premiss together with the other premiss that bodily identity and memory continuity are the only two criteria that are there for personal identity.

as Fawkes. But it cannot be both that Charles is identical with Fawkes and that Robert is. For in that case Charles and Robert, two contemporary (and hence different) persons, would be identical, and this is absurd. This absurdity could be avoided (i.e., one could try to avoid it) by abandoning one or both of the assertions: "Charles is Fawkes" and "Robert is Fawkes". Williams argued that it would be "vacuous" to assert one of these assertions and abandon the other, since, ex hypothesi, there would be nothing to choose between them; hence the natural course would be to abandon both. And therefore, he argued, it would be as vacuous to make identity judgment when Charles alone undergoes the change as it is in the reduplication case. It follows from this that memory-continuity (in the relevant sense outlined in the beginning) cannot be a sufficient condition of personal identity. Further, that it cannot be a criterion of identity in the first place is pleaded by Williams by invoking an important principle which a supposed criterion of identity must satisfy. The principle is that identity is a logically one-one relation, and so, that no principle can be a criterion of identity if it relies on a relation that is not logically one-one. (A relation is logically one-one if and only if it can relate an earlier  $p$  to one and only one later  $p$  - whatever  $p$  may stand for). It is evident that the memory criterion of personal identity does not satisfy this requirement. For Williams' argument has made it clear that 'making the same memory-claims and/or having similar character and personality as' is a relation that is logically many-one, and that in this respect many contemporaneous persons can claim to be identical with one particular person at the same time.

Williams' argument has since remained a substantial setback for any form of memory-theory; and the recent talk of "non-branching psychological relations"<sup>3</sup> seems to be only desperate attempt at salvaging it. The fact is that any theory that leaves room for reduplication must be abandoned; and

the truth is that the memory criterion does leave such room. (It will be argued later on that "non-branching psychological relation" theories do not quite succeed in ruling this out).

Besides proving the deficiency of the memory criterion (and that even as a criterion of identity), Williams also wanted to show (if the latter did not show, or was not enough to show, this) that bodily continuity (which he specified, includes spatio-temporal continuity) is a necessary condition of personal identity. For the principle of identity, or rather the test of an identity-criterion, which he formulated and which the ~~memory~~-criterion failed to satisfy, accommodated the bodily criterion beyond any reasonable doubt. "Having the same body as" is a relation which is logically one-one and cannot, without contradiction, relate more than one person (at any particular time) to an earlier person. Moreover, it is only in respect of bodily continuity that one can distinguish between identity and exact similarity in the case of persons. We can easily say which two (non-contemporaneous) persons are identical (namely, those who have the same body) and which persons are exactly similar, but not identical (those having different bodies). Yet this we could not say if similarity of memory-claims and of personal characteristics were our criterion of personal identity. For not only that one and the same person can display similarity of memory etc, but two (or twenty or two thousand) different persons can also display the same - even at the very same time - as is evident from Williams' case of Charles-Robert/Fawkes. The fact is that without spatio-temporal continuity (which cannot intelligibly apply in the case of memory and character in absolute exclusion of the body) the notion of identity or sameness is not intelligible; and if we do speak of same memory we can only mean "exactly similar memory" by it for that is what it means. But, as Williams rightly claimed, "same body" and "exactly similar body" really do make a difference.<sup>4</sup> And following this,

one can say that 'same body' does point to the same person as distinct from an exactly similar person, whereas there can strictly be no 'same memory' to point to the same person: all we are entitled in a memory theory, is only to talk of persons at different times that they are exactly similar, but of no two of whom can we have any ground or justification for saying that they are the same.

The conclusion we are expected to draw from these considerations is that memory-continuity cannot be a sufficient condition of personal identity and hence that bodily continuity is a necessary condition of the latter. Nevertheless Williams does not rule out entirely the possibility that a later person should be identified with an earlier person without reference to a body. But, he claimed, "it is a necessary condition of making the supposed identification on non-bodily grounds (i.e. on the basis solely of the memory-criterion) that at some stage identification should be made on bodily grounds".<sup>5</sup> This is presumably to harp upon the well-known claim that the memory-criterion is only a secondary criterion of personal identity and is dependent on the continuity of body of the person whose identity is in question.\* Williams' argument for this seems to be that if a later person ( $p_2$ ) claims to remember what an earlier person ( $p_1$ ) had done, identification can be made by checking if  $p_1$  had done what  $p_2$  claims to have done; and this checking is possible only by reference to witnesses of  $p_1$ 's activities, and these witnesses must have seen  $p_1$ 's body the continuity of which must be relied on "in order for their accounts to be connected into the history of one person."<sup>6</sup> What this means is perhaps that the witness who had observed  $p_1$ 's body being involved in the remembered action must presume that this body is spatio-temporally continuous with the

---

\*The nature and force of this argument will be brought out more elaborately in the next section and also in Ch 4 sec iii.

body of  $p_2$  in order that the latter may be said to remember (in the strong sense of that word) having done the action: if that presumption is known to be false (as in the case of Charles/Guy Fawkes) then  $p_2$ 's memory-claim would turn out to be merely apparent, and judgment of identity withdrawn. This is not to beg any question. For the argument here is not  $p_2$  remembers  $x$  (what  $p_1$  did), therefore  $p_2$  is the same person as  $p_1$  (who did  $x$ ) but rather:  $p_2$  claims to remember  $x$ , it is reasonable to believe that he is remembering, therefore it is reasonable to believe (or say) that  $p_2$  is the same person as  $p_1$  (This is all we are entitled to say, for a criterion never entails the truth of what it is a criterion of, see last section). Speaking generally, it is only because there is such a thing as real memory that memory-criterion (memory-claims, not real remembering, see Williams Ibid p 4) is used as a criterion of personal identity at all, i.e., a memory-claim creates the presumption that it may be real remembering and so points reasonably to personal identity. But a real remembering there can be only if the person 'remembering' an action or an event was physically present to do the action or to witness the event in question, and this latter could be checked only by observing the continuity of the body (or, at least the brain, as we will argue later on) that was involved in the **remembered** act or event, up to the time when the action is remembered. If this is correct, then Williams rightly says that any claim that bodily considerations could be absolutely eliminated from the criteria of personal identity must fail.

One of the most perplexing counter-examples to the bodily identity theory are the putative cases of bodily interchange or bodily transfer. Williams is found to come to grips with such alleged possibilities at several places.<sup>7</sup> But nowhere he seems to have been as successful on this score as he generally has been in his attack against memory-claims being the sufficient condition of personal identity. As a matter of fact one rather gets the impression that he has given a much larger concession to

the change-of-body hypothesis in his otherwise prominent paper "The Self and the Future", rather than providing any very strong argument against it. In general, Williams' arguments against this possibility seems to have been two-fold. One aspect of his argument seems to draw force from an appeal to our logical intuitions: although certain situations could be so conceived as to earn the "change of body" description up to a certain extent, still there are what he calls "logical limits" to the thinkability of such situations. A second aspect of Williams' arguments relates to what the original persons can expect with regard to their own future; though on the face of it the change-of-body hypothesis is not to be entirely ruled out, yet it cannot be taken for granted and that, being faced with a proposal to undergo such a process (of alleged change of body), each "original" person would perhaps like to take the "risk"<sup>8</sup> of identifying his future self with his present body and would accordingly choose all good things to happen to the person with that body (for as Williams took it, the theory may or may not be right). Both these arguments seem content to expose the supposed hypothesis as rather dubious than proving it to be impossible or implausible in any way. I want to suggest that even as such attempts they are not quite successful. We must now consider the arguments one after another.

The first argument allows it to be possible that two persons, an emperor and a peasant, say, may (or may be made to) have each other's memories, character and personality etc; but it finds difficulty in the possible display of the latter. "How could the peasant's gruff blasphemies be uttered in the emperor's cultivated tones?", Williams asks. And similar considerations go for features like facial expressions and characteristic smiles (see Ibid p 12). Neither the emperor's face could express the "morose suspiciousness" of the peasant, Williams claims, nor could the peasant (his face) wear the "characteristic smile" of the

emperor. Consequently, neither could be the same sort of person as the other and, so it is argued, bodily interchange cannot be taken for granted. Now, the first thing that suggests itself is that none of these limitations is "logical" in any sense. The above situations may be unthinkable (if it means unimaginable by some or many or most), but not (logically) impossible - or, perhaps they are not even unthinkable in the above sense (for can't many imagine them to happen?). There are cases of plastic surgery by which these difficulties (which, pace Williams, are merely empirical) could be overcome. Moreover if actings and mimics are often quite accurate, to the point of being delusive, there are no logical difficulties as to why they cannot be perfectly so as to look as if they are not acting but indeed personality change (for is this not what would suffice for Williams)? The second thing about it is that even if Williams is right about the limitations one still fails to see how far they count against the bodily-exchange hypothesis. It rather seems plausible that these limitations need not disturb the proponent of the change of body hypothesis, since he is a memory theorist for whom the continuity of bodily features is utterly inessential. All that matters for him is that the peasant-body-person has the memories and character of the (sometime) emperor and vice versa; and this Williams would apparently allow to be possible.\* And as regards the exchange of character and personality, the satisfaction of this requirement<sup>1</sup> is not necessarily hindered by the lack of the relevant physical features. Would it not be enough, for example,

---

\*It is evident from the fact, indicated earlier, that Williams made greater allowance for this in his "The Self and the Future".

<sup>1</sup>The satisfaction of this requirement, though, is not necessary to a memory theory of personal identity. As noted in the beginning, similarity of memory claims may or may not be accompanied by that of character and/or personality etc.

if the present emperor-body person displays a natural skill in farming and harvest prediction and the present peasant-body-person shows acute interest in politics and governance? Moreover, even to a bodily theorist, continuity of physical features like characteristic smiles and grimaces are not essential to personal identity. If I survive a fire and retain a badly burnt face, I cannot smile the characteristic smile I used to; but that does not make me a different person (I certainly do not expect any bodily theorist to say that). Finally, one does not (in these respects or in any other) have to be the same sort of person in order to be the same person. What one may become is a changed personality - which admittedly is a different matter.

The second point in Williams' argument is brought out by presenting two alternative descriptions of a thought experiment which is designed to effect a mutual transfer of memories and character etc, between two persons without letting any part of their bodies to be altered or exchanged. (Hence the plea for a special device of information - transfer from their brains).<sup>9</sup> If the result of the experiment is complete reversal of the personal memories and characteristics with striking accuracy, as in fact is shown by Williams first picture of the experiment, then the "change-of-body" description of the situation might seem quite plausible. i.e. if the original persons are A and B and if after the experiment the A-body-person displays all the "B-ish" characters and makes memory claims that fit the life of B\* and vice versa, then A and B might plausibly be said to have exchanged bodies. But, indeed, this cannot so surely be the correct description of the situation - particularly if one of the persons to be experimented upon is oneself. This Williams tried to show in his second

---

\*It is to be understood also that this person does not display any of the characteristics and memory-claims that formerly went with A (and vice versa).

picture of the situation. This picture, which we have called an alternative description of the thought experiment, purportedly exploits the incompatibility - or, at least the indeterminacy - of the relation between one's expectations with regard to one's future <sup>of</sup> and the change-of-body hypothesis. The expectation considered here is that of future (bodily) torture or, simply, pain. Williams' purpose here is to emphasize the fact that if physical torture is going to be inflicted on my present body then I am going to feel the pain, and nothing about my future psychological condition will remove or reduce my feeling of fear - not the assurance that I will not remember having been told about it, nor that I will not remember anything about my past and that, before the infliction of the torture I will have a completely different set of memories, nor even that I will acquire the "memories" of someone else who, by a similar operation, will have "my" memories. (The quotations in this sentence are used to present as neutral a description of the situation as possible, and not to beg any question). On the contrary, it is argued, my feeling of fear will be only compounded at the prospect of suffering from mental derangements and unexpected torture. It seems to follow, therefore, that despite the alleged prospects of a complete psychological change or even exchange with someone else, one still will (is likely to) be concerned over what is going to happen to one's present body (or to the person with that body). And this means that, given this description of the situation, the bodily-exchange hypothesis cannot be taken any seriously, and that mental identity cannot be so sure a guide to personal identity as bodily identity is.

Thus, Williams gave two alternative descriptions of the thought experiment designed to effect memory-and-character exchange; and whereas the first description, carried in third-personal terms, pointed to the mental identity criterion (bodily-exchange) the second one - carried in

first personal terms - pointed to the bodily identity criterion of personal identity (this fact has been pinpointed by Williams on p 62 of Problems of the Self). But the experiment, in order that it may establish a reliable theory, must not lead to different results - when carried upon others and upon me. For it is not only the case that personal identity means the same in both cases, it is also not desirable that the criteria of personal identity should be different depending on whether or not the person is oneself. It seems reasonable, therefore, that Williams should be reluctant to accept the 'change-of-body' hypothesis despite its seeming so plausibly to follow from his first picture. As he claimed, there seems to be nothing wrong with his second description of the thought experiment so that we may be persuaded to accept the first in preference to the second. Without in fact much bothering to argue for the acceptability of his second description of the experiment, a bodily theorist as he is, Williams gives his preference for it simply on the ground that it needs to be shown what is wrong with it (Id p 63) - though admitting that his choice is "risky".

Or, perhaps, he did have some arguments to substantiate his preference. Earlier in the same paper, he had claimed that "memory is a causal notion, and ... it seems a necessary condition of x's present knowledge of x's earlier experiences constituting memory of these experiences that the causal chain linking the experiences and the knowledge should not run outside x's body".<sup>10</sup> and, that our thought experiment did not provide for this condition to be satisfied (cf Ibid p 56). But without this condition being satisfied, my "new" memories (induced in me by means of the experiment) can hardly count as memories of a different person, and I would be clumsily hovering over the borderline between being the person with that other body and being merely a clairvoyant, but nevertheless the same person as the one with my present body. This is the element of "risk" that is involved. And once again, <sup>the</sup> question

of identity and exact similarity would return: the new set of memories induced in me may at best make me a person exactly similar (in that respect) to my coparticipant in the experiment; but surely the obvious discontinuity (non-identity) of our bodies will show that I am not him. And since the choice here is either that I will be the same person as the person with the other body or that I will be the same person as the person with my present body, it would surely be reasonable to think that the latter is the case, even after the experiment has been successfully performed. Thus the change-of-body proposal is ruled out and this would explain Williams' preference for the second description.

It will be said that the change-of-body proposal is not logically ruled out since the supposed ground of its dismissal is not logically necessary. In other words, it will be said that it is not a necessary truth that memory is a casual notion and more particularly that the casual chain involved in it must run through the rememberer's body. This is true and most widely agreed. But from what has just been said above it will appear that this principle, though logically contingent in itself, is indispensable to the possibility and reasonableness of an important sphere of discourse. This is what may be called a 'transcendental argument',<sup>11</sup> in favour of the said principle. If the latter is not accepted as true, we cannot possibly tell memory from apparent memory or mere seeming memory. The last distinction is a conceptual distinction, and this will be obliterated if the two concepts of memory and seeming memory could not be possibly differentiated in their application to the world. But as have just been seen above, without a casual explanation of the said type we will have no justification whatsoever for making any such differentiation. Further, in view of the logical connection between memory and personal identity, memory would cease to be a criterion of the latter, since it could never justify anyone in saying that a later person is the same

person as an earlier person. For no knowledge that people remember can ever be obtained. This will, thus, be too high a price to be paid if we are not to accept the described casual principle of memory. And if this is so, the alleged change-of-body hypothesis will have a double disadvantage since, the only criterion of bodily continuity being inapplicable here, it cannot be put forward as a reasonable hypothesis of personal identity (or non-identity).

Now, of the two arguments Williams put forward against the change-of-body hypothesis the second, outlined in the paragraph before the last, is more convincing and less vulnerable than the first, and that it can be more reasonably substantiated to show that the change-of-body hypothesis is at least highly dubious. Yet for the success of this argument Williams seems to rely heavily on a principle which, even by some of his own admissions, does not seem to be so sound as one would wish it to be. The principle is: "my undergoing physical pain in the future is not excluded by any psychological state I may be in at the time" (Problems of the Self, p 53). Were it that this was always true! Williams himself envisaged possible exceptions to this principle (and consequently, we should add, to his second picture of the experiment): Suppose A suffers from acrophobia and so, when he is told that he is going to be thrown off a steep mountain he is horrified. But if he is told that before being thrown he is going to be completely cured of acrophobia then, Williams agrees, he will not fear the fall or at least "will not have the same grounds of fear". And to this he promptly adds, "physical pain ... is absolutely minimally dependent on character or belief. No amount of change in my character or my belief would seem to affect substantially (my emphasis) the nastiness of torture applied to me" (p 54). But if the intensity of fear would be still affected to some extent by the change of my character etc, why could it not be possible for it to be totally affected? In the above case,

acrophobia was the cause of the 'fear of the fall from a mountain', and it was eradicated by the eradication of acrophobia. But, Williams is content to point out, there would still be fear - though not the same ground for it. But can we not think of the possibility that psychologists could trace fear (the general feeling) to some sort of psychological condition so that by the eradication of that condition one could be completely cured of fear? That this is thinkable and, hence, logically possible can be a fair reason against Williams' thesis about physical torture and fear for that and yet a lot seems to depend on that thesis in order for Williams' second argument against the change-of-body hypothesis to succeed. However, the more important question, I think, is not whether I am going to feel the pain (that is going to be inflicted upon my present body in the future), but ~~whether~~ the person in this body will be a different person if I don't. And although the answer to the former question may be indeterminate or possibly "no", the answer to the latter is quite certainly "NO". People are often anaesthetised or they become mentally deranged, but for that reason they don't cease to be the persons they used to be nor do they become different persons. What prevents this latter consequence from being deduced is the continuity of their bodies; and therefore indeed it seems that whether or not my reaction to the experimenter's proposal is fear (and accordingly, whether or not I think that I am going to feel the pain) this has very little to do - if at all - with the present-bodied-person's being the same person as me or a different one after the experiment. (Williams' long drawn argument in his second description of the experiment - designed to neutralise the force of a 'change-of-body' picture made plausible by his first - seems hardly to have served his purpose. If he saw "nothing wrong" with his second description our immediately preceeding analysis should have indicated that something is wrong). All that is needed is a rather straightforward demonstration

that bodily continuity is necessary condition for personal identity which would definitely deny the bodily interchange hypothesis rather than making it just dubious; and I do not think Williams has ever succeeded in doing this. In what follows, I shall try to make such an idea plausible by utilising elements drawn from Williams' arguments. In this attempt, I shall be making a slight modification in the concept of 'bodily continuity' which, however, I shall claim to be quite compatible with Williams' basic commitments. More particularly, I shall try to show that the puzzle-situations that suggest bodily transfer (and so threaten the necessity of bodily identity) are either implausible or, if not, are such that one can always explain them in terms of bodily continuity in the particular sense which I shall be giving to the latter.

Now there can be two\* types of puzzle situations suggesting that a bodily transfer has taken place. The first type, which owes its traditional origin to Locke<sup>12</sup> and which in recent times has been contemplated by Quinton,<sup>13</sup> among others, works upon the mere possibility of an apparent change-of-body evidenced by the accurate display of memory-claims and character etc of the first person in what used to be the second person's body and vice versa - "mere" possibility, because the proponent of such hypothesis does not say how this alleged change has taken place (let us call this the "naive thesis" of change-of-body). The second type of situations suggesting a bodily interchange does have an explanation as to how it has come about, namely, that the brains of the two persons have been inter-transplanted (let us call it the "scientific thesis" of change-of-body). This type of cases have been advocated by Shoemaker, among others, in his Self Knowledge and Self Identity (pp 23-24).

---

\*A third can also be contemplated, but can be more easily got out of our way. We shall come to that a little later.

Any proposal to describe these cases as cases of bodily transfer carries with it the implication that the memory claims made by the (later) persons concerned are genuine memories, for it is solely this presumption that gives sense to the change-of-body hypotheses in the first place. But we have already seen the importance of the fact that genuine memories must satisfy a casual condition which requires that the casual chain linking the present knowledge ("memory") and the earlier experience (that is being "remembered") must not run outside the body of the rememberer. But ex hypothesi, with two numerically distinct bodies in the case of each "rememberer" and with no casual link whatsoever between these bodies, the Lockean-Quintonian type of cases cannot satisfy this requirement for memory claims being real memories. It is on this ground, as we saw, Williams would dismiss <sup>the</sup> "naive thesis" as implausible. This condition, however, is satisfied by what we have called the "scientific thesis" which, only with that motive, introduced the inter-transplantation of brains. Of course Williams will still not call them 'clear cases' of bodily transfer.<sup>14</sup> What I shall try to suggest is that, call them by any name you like, change-of-body or otherwise, the important fact is that the plausibility of these alleged cases is always compatible with the necessity of bodily continuity and, therefore, that the purported suggestion that all these possibilities show the memory criterion to be the sole or even the primary criterion of personal identity is just mistaken.

But before I proceed to do this I will mention a third type of possibility (apart from the two accepted for consideration) that may suggest a bodily transfer and the consequent primacy of the memory criterion. This is being mentioned not to be pursued any further but simply to get it out of our way. This type of possibility suggests that memory and character may be transferred from one person to another - not by brain transfer nor by the transfer of any bodily part, but by means of a

transfer of information between brains. Such a possibility was contemplated by Williams himself - somewhat vaguely in his "The Self and the Future" and quite elaborately in "Are Persons Bodies"? It is imagined that all the information may be taken out of and kept in a special storage device, and then put back into my brain (say after cleaning or repairing the latter): it is also imagined that instead of the information being put back into my brain it might as well be put into someone else's brain with the result (it is supposed) that out of (what normally would be called) the other person's lips came the memory-claims that are not "his" but "my" memories. (Similar operation can simultaneously be carried out in the reverse direction so that "his" memories would come out of "my" lips). Now the merit of this type of case as cases of alleged bodily exchange seems to be that it purports to overcome the difficulty concerning the casual requirement of memory discussed above. For, As Williams admitted "... at least we can be clear that passage of information via the device is not in itself incompatible with the later knowledge's being memory."<sup>15</sup> But here again the price to be paid for this advantage would be too high; for there will, in that case, be no guarantee against the possibility that the information from the 'special device' may not be passed on to many different brains housed in the bodies of many persons, and because of this possibility of reduplication there will not be identity in the first place. (This, incidentally, shows that for the purpose of personal identity the satisfaction of the casual condition of memory, though necessary, is not sufficient.) Thus, the third type of bodily-interchange hypothesis fares no better than the first in establishing personal identity on the basis of memory claims.

Back now to the Shoemaker-type cases, which we have called the 'scientific thesis' and which not only satisfies the casual requirement

of memory but also is immune from the possibility of reduplication.\* For if we agree to say in such a case that B is (or has now become) the same person as A, we would do so on the ground that B displays the same memory- and character etc as did A because B's body contains now the same brain as A's body did; and further it cannot be the case that there be two bodies (or more), each containing precisely the same brain (A's brain) at the same time.<sup>16</sup> In such a case, Williams says, the judgments of personal identity "might reasonably" go the way of the character-and-memory traits and if so, he adds "we would here have a divergence from bodily identity".<sup>17</sup> But, as promised earlier, I shall now try to suggest, on what I shall claim to be essentially Williams-type grounds, that this description of the case is compatible with the necessity of the bodily identity criterion. To this purpose, I shall approach the case with its exact way of presentation and <sup>an</sup>alysis as was originally offered by Shoemaker: Two persons, Robinson and Brown, have undergone a brain operation and by mistake (after the necessary cleaning of the skulls etc), Brown's brain has been placed in Robinson's skull and the latter's brain in the former's skull. One of them immediately dies and the other, with Brown's brain and Robinson's body, eventually survives; Shoemaker calls this person Brownson. Brownson makes all the memory claims and displays character and personality etc that only the late Brown could have done; he recognises the dead body as his body and disowns the body which, as it were, he now "inhabits". This case, if it could come to pass at all, has been designed to show that Brownson is really Brown and hence that personal identity is based on the mental criteria - there being nothing of bodily continuity excepting the continuity of (Brown's) brain. But then, in the event of such a situation

---

<sup>it</sup>  
\*So it will be if/takes, as it does, the whole brain to be the bearer of personal identity. In case of split-brain transplant the difficulty of reduplication would reappear, and I have further reasons to say that the latter type cases will not be reasonable cases of change-of-body (see next section).

occurring, I do not see anything "absurd" (pace Shoemaker<sup>18</sup>) in saying that brain continuity is our criterion of personal identity here. It is an established fact that the brain plays the central role in the "essential and characteristic vital functions" of a person, and it would be only fair where in strange but plausible cases like this, neither the bodily continuity (continuity of the whole body) criterion nor the memory-continuity criterion seems satisfactory, to take the continuity of the brain as the "individuating nucleus". And if this explanation is right and acceptable, bodily continuity (i.e. continuity of the brain) would still be a necessary condition of personal identity. Shoemaker rejects this explanation as absurd on the following ground: if the outcome of his brain-transfer operation were different and if Brownson were to act and talk like Robinson "surely none would say that this man who looks and acts and talks just like Robinson and has what has always been Robinson's body, must really be Brown rather than Robinson because he has Brown's brain".<sup>19</sup> True, no one would or should call this new person Brown, in this changed situation; but why can't we call him Robinson now? For to be sure, here we have not only bodily continuity (in the more familiar sense) but also psychological continuity to our advantage! Presumably what for Shoemaker would count against our calling him Robinson is the lack of the close casual relationship that is supposed to hold between the state of a man's brain and his psychological features. But why should this be so puzzling in a puzzle situation like this? After all, as Shoemaker himself saw, "whatever relationship there is between the state of one's brain and the state of one's mind (i.e. one's psychological features) is surely casual and contingent, not logically necessary",<sup>20</sup> and hence it is fairly conceivable (logically possible) for the relationship not to hold. Moreover, if it is conceivable (logically possible) for a man (for this man whom we propose to call Robinson) to display the same psychological features as he used to have even without his own brain and with somebody else's brain, this

would be a clear case of the usual (casual) relationships not holding. After this nothing should prevent us from saying that Brownson is the same person as Robinson. Of course this would still leave us with the difficulty of explaining what, if not the brain, is then the seat of memory and consciousness etc under this situation. I do not know the answer, and perhaps nobody does, in a single case like this. But let us try to generalise the case and imagine that both the post-operative persons in Shoemaker's example survive and behave as usual" i.e. Brown (with Robinson's brain) as Brown and Robinson (with Brown's brain) as Robinson. Let us also imagine a world where these things are of frequent occurrence so that they are no longer to be taken as puzzle cases in this world. Certainly in such a world our existing hypothesis that brain is the seat of memory and other mental capacities (let us call them all "personal faculties") would be just false. Perhaps it would be reasonable to suppose that in such a world, the seat of the personal faculties would be somewhere in the human body and that the brain (any brain) acts simply as a catalyst, as it were, to activate it so that only with a brain (and not without any) a person may be able to exercise his personal faculties. Admittedly, such a description would be upsetting to our currently accepted theories of mind-brain relationships and of personal identity; but why should we try to apply our standards in this type of a world which is not recognizably ours? If we are certainly entertaining the possibility of a different world, we should "bear in mind the notion of open texture and the Wittgensteinean idea that how a given occurrence or state of affairs should be described depends not on the occurrence or the state but also on surrounding circumstances."<sup>21</sup> Thus, if what is said in the foregoing is right, it is shown that bodily continuity as a necessary condition of personal identity can very well be said to obtain also in puzzle cases like Shoemaker's and in my amended version of it (both of which will fall under what we have called the "scientific thesis" of alleged change-of-body).

I am fully conscious that while giving my above account of bodily continuity I have been using the latter in quite a different sense - not particularly in the sense in which Williams wanted to make it a necessary condition of personal identity. More particularly, I have been taking the continuity of the brain (and for that matter, of any bodily part) - not necessarily of the whole body\* - as bodily continuity. My plea for doing this is to strengthen Williams' arguments - to present what I think to be a more consistent and persuasive form of Williams-type arguments. For I think that it is the basic purpose of Williams in arguing for the necessity of bodily continuity criterion that it (bodily continuity) involves spatio-temporal continuity. That Williams emphasizes spatio-temporal continuity for personal identity has been hinted at earlier and comes out most expressedly in the following passage, where he sets out to argue that since spatio-temporal continuity is "interfered with" in the case of fission,<sup>22</sup> judgments of identity should be withheld in such cases. Outlining an apparent objection (issuing from the fission case) to his criterion of spatio-temporal continuity he says:

"... it may be said ... that even a criterion of identity in terms of spatio-temporal continuity on which I lay the weight for personal identity is itself not immune to this possibility (of redeuplication PKM)"<sup>23</sup> (my emphasis).

It would seem quite plausible, therefore, that for Williams, what is important for personal identity is spatio-temporal continuity and that if for him, bodily continuity and personal identity must go hand in hand, this is because, under all normal circumstances, spatio-temporal continuity of a person can apply (or is applicable) only to the continuity of his body. In view of this, our modification in the concept of 'bodily

---

\*I have stronger arguments to show that the idea of 'whole body' cannot be maintained and is not necessary to personal identity - all this in the next section.

continuity' will not impugn Williams' scheme of things, since continuity of the brain serves basically the same purpose which the continuity of the whole body was designed to serve: preserving spatio-temporal continuity.\* This should rather be a welcome procedure, for in certain abnormal cases like what we have called the 'scientific thesis', where it might be reasonable (as Williams said it is) to say that personal identity should go the way of memory-and-character, identity can readily be explained in terms of bodily (brain, or any other bodily part that played the same causal role as the brain) continuity and not in terms of memory-and-character continuity - much less in terms of the latter solely or even primarily. If it can be so expressed, memory-continuity in such cases, would be the effect, not the cause, of personal identity, whereas bodily continuity (in our modified sense) would be the cause of personal identity. Or, if you like, memory continuity would be due to personal identity while personal identity would be due to bodily continuity.

Given this modified version of bodily continuity, a Williams-type theory of personal identity would run roughly as follows: spatio-temporal continuity of the body, or rather the part thereof which is causally responsible for the 'personal faculties' (memory, character, personality etc) is always a necessary condition of personal identity; similarity of memory and character-traits alone and unsupported by any casual explanation is not even a sufficient condition of personal identity.

This theory will have the merit of explaining all the normal cases of personal identity; it will also explain the abnormal cases where judgment of personal identity has reasonableness and the relevant causal support.

The theory will have the following consequences - consequences which

---

\*This at least is a necessary condition, though not sufficient; but of that later.

Williams is obliged to entertain, yet which, because of his limited notion of bodily continuity, Williams failed to embrace: change-of-body will be ruled out as logically impossible. If a situation could be so created viz, by the physical transfer of brains, "change-of-body-minus-brain" would of course be the only correct description of the situation; but this would NOT mean a divergence from the bodily identity-criterion since something importantly\* bodily is still continued. Nor - and more importantly - will it show (what the proponents of the bodily transfer hypothesis claimed it shows) that memory-and-character continuity is the only or at least the primary criterion of personal identity.

Thus presented, Williams' position - and for that matter any correct bodily identity theory - is clear and straightforward. Yet there is one very recent attack against Williams which must be considered; the effect of considering this will, I believe, show not only the strength of Williams' position but also a successful application of the principle underlying our account of the bodily continuity theory. This attack has been levelled by Professor Vesey in his book Personal Identity. The attack is significant not because of any merit it has but because of its claim to have utilised the causal aspect of memory and other mental capacities, which we have seen to be so important for Williams' theory and for our (modified) account of bodily continuity. But we shall see that the attack is entirely mistaken and that the idea of the causal (or as Vesey calls it, 'physical') basis of memory etc, which Vesey thought would work against Williams, does indeed defeat his objection and

---

\*Importantly" from the standpoint of personal identity. The rationale of this qualification has been suggested reasonably well in this section and something more of this is said in the earlier pages of the next section.

expose its barrenness.\* Vesey used his attack pointedly against Williams' argument, adduced in his famous case of Charles - Robert / Guy Fawkes, that similarity of memory claims etc, cannot be a sufficient condition of personal identity. After arguing that neither of the claimants can be identical with Fawkes, Williams concludes:

"so it would be best, if anything, to say that both had mysteriously become like Guy Fawkes, clairvoyantly knew about him, or something like this."<sup>24</sup>

And (this is the gist of his argument)

"If this would be the best description of each of the two, why would it not be the best description of Charles if Charles alone were changed?"<sup>24</sup>

Vesey's first complaint against Williams seems to be that his case is "far-fetched" and, as such, obscures the understanding of what it proves - though it does prove "something".<sup>25</sup> It is far fetched because, Vesey thinks, it has no basis in what we know about the physical basis of memory. To make the case a little "less far fetched" Vesey imagines cases of bisected brain transplants (originally devised by Wiggins and Shoemaker) and applies it to Williams' case by supposing that Guy Fawkes' brain could be bisected and preserved after his execution and, later on, each half could go to what used to be Charles' and Robert's bodies - with the result that out of the lips of "Charles" and "Roberts" came the memory claims that fit the pattern of Guy Fawkes' life.

Before going to examine what Vesey thought to be the implication of this modified case on Williams' argument, a word about "far fetchedness". If Williams' original case was far fetched, in what way Vesey's case of bisected brain transplant is not? After all Vesey himself was once obliged to describe it as "rather fanciful" (Ibid, p 85)! Then is a case

---

\*Since it is precisely the lack of this justification or one like it that made Williams say what he did, and if one such justification is provided, he would not say that; - consequently the objection would be pointless.

being "far fetched" only a matter of degree? If so, where to draw the line? Perhaps the brain transplant case is not as hopeless as Williams' case is, and this is probably because the former can be practically possible (cf Vesey's "post script" on p 86) and would thus provide a physical basis of memory. Granted that this is so, would it provide any logical guarantee that the brain is the physical basis of memory? On the contrary. It will be recalled that we showed earlier that the seat of memory might be somewhere else in the body and a brain (any brain) might be only instrumental in activating it - so that it is fairly possible that despite having Fawke's brain fitted to his body Charles may make memory claims that fit his own life rather than that of Guy Fawkes. But if this is possible, it would disprove or at least undermine the purpose of the claim that brain is the physical basis of memory and other mental capacities, and hence Vesey's case of (bisected) brain transplant would lose its purpose. The fact is that a case being far-fetched is never a disadvantage of its being an instrument of philosophical arguments. As long as it is logically possible, it will do the trick; and so does Williams' original case.

Now, about Vesey's modified version of Williams' story and what impact he thinks it may have on the latter's argument. Vesey thinks that this modification will present Williams' question as follows:

"If there were a bisected brain transplant ... it would be absurd to say that the two people were both Guy Fawkes since it is absurd to say that Guy Fawkes is in two places at once and absurd to say that two people were identical with each other. Further, there would be no reason for saying that one, rather than the other, was Guy Fawkes. So if there were whole brain transplant (all of Guy Fawkes' brain going into what used to be Charles' body) would it not be vacuous to say that Charles is now Guy Fawkes?"<sup>26</sup>

And he asks, why should it be vacuous? My immediate reaction is that Vesey here is only shooting at a Strawman. For Williams never thought to - or at any rate never did - apply his argument by means of such type of cases, and when the situation is thus altered one doubts (and we have given

enough reason to doubt) whether he would still like to argue in the way Vesey thinks he would. Particularly, I do not think that he would say it is "vacuous" to say that Charles (with the whole of Guy Fawkes' brain) is now Guy Fawkes. For the reason why (in his original case) he said that it would be "quite vacuous" to say that Charles is Guy Fawkes even if Charles alone were changed is that there would be no "grounds"<sup>27</sup> for such identification; but here with the whole of Guy Fawkes' brain transplanted into Charles' body there would be a very strong ground for such identification. Moreover, as I have shown, making identification in this case would not impugn Williams' claim that bodily identity and personal identity go hand in hand since strictly preserved here is spatio-temporal continuity (of something bodily - the brain) of which bodily identity is a specific case. And as regards bisected brain transplant case, it may not be "vacuous" to say that both the claimants are identical with Guy Fawkes, but saying that will take us from frying pan to fire, for firstly it will do violence to a very fundamental requirement of identity (which Williams so much stressed), namely that it is a logically one-one relation - the requirement which, incidentally, would secure the claim of whole-brain as the bearer of personal identity. (No two contemporaneous persons can be coherently said to have the whole brain of a third person). And secondly, as Williams would argue,<sup>28</sup> this case, being a fission-like case, would "interfere with" spatio-temporal continuity and therefore would force us to answer the identity question in the negative.

Vesey's attack appears more ineffective and unattractive when he goes on to answer Williams' question by taking the help from Parfit's suggestions. Williams had argued that identity is a logically one-one relation and that "similarity of character and memory claims", if taken as sufficient condition of personal identity, would leave room for reduplication. Parfit tried to answer to this by suggesting that psychological

continuity (which Vesey and, dubiously, Parfit take to be indistinct from similarity of memory claims) can serve as a criterion of identity when it is one i.e. by claiming that "non-branching" psychological continuity is logically one-one.<sup>29</sup> Taking this hint from Parfit, Vesey suggests an answer to Williams' question by saying that Charles is identical with Guy Fawkes, "Because in the case of Charles alone being changed the psychological continuity is non-branching".<sup>30</sup> This argument is extremely unattractive because Williams never used his argument - and never meant to use it - against the criterion of "psychological continuity" but against that of similarity<sup>of</sup>/character and memory claims; and these two are not quite the same. I shall try to explain the difference.

"Psychological continuity", as Parfit used the expression, is a technical expression which is not unsupported as mere similarity of memory claims is. On the contrary, as far as I have tried to understand it, psychological continuity carries with it a causal support and indeed the support provided by the human brain. Parfit, who invented the expression as well as the idea, never used it - except in few of his unguarded moments, like the one under reference - to mean the same thing as mere unsupported memory claims. (Though it is rather strange that he himself takes the help of this concept to reply to Williams). We have ample evidence that he used this expression in connection with fission, fusion and brain-transplants, all of which preserve spatio-temporal continuity - including causal continuity. That he used the expression more often in connection with brain transplants is evident from Vesey's own illustrations of Parfit's standpoint (see Personal Identity, pp 88-89). And that the expression was used by its author to be a causal notion can be evident from the following remarks of Parfit:

"I shall assume that, after I die, God will create a replica of me ... Between me and the replica what relations holds? The answer is: psychological continuity with a special cause. The normal cause is (we believe) the continuity of the brain".<sup>31</sup> (my emphasis)

Again:

"... Nor does it matter in the slightest that the psychological continuity will lack its normal cause. All that is needed is a reliable cause. ... This reaction seems to me wholly reasonable." <sup>31</sup> (my emphasis)

All this should be enough to make it clear that Williams never meant to apply his argument against the supposed criterion of psychological continuity in the technical sense in which Parfit and, following him, Vesey used it. As far as I have tried to find out, Williams never even used the expression "psychological continuity" in that sense. His attack was rightly directed against the criterion of "similarity of character and memory claims" which it should by now be clear, is not the same as "psychological continuity". Now, given the case of the whole brain transplant and the consequent psychological continuity, Williams, I am quite sure, will not ask the question he asked about his Charles/Guy Fawkes case; and moreover for the reasons we have given earlier, he probably would - or rather should - agree to call these cases, cases of identity.

It follows, therefore, that Vesey's answer (and also Parfit's, in so far as it is the same) to the Williams' question is an answer to a question that does not exist. And added to this, Vesey is certainly wrong in accusing that Williams, unlike Parfit, thinks that there is more to personal identity than non-branching psychological continuity. <sup>32</sup> For in view of the distinction we have drawn between "psychological continuity" and "similarity of memory claims", a more just and sympathetic reading of Williams would be: There is certainly more to personal identity than mere unsupported similarity of character and memory claims.

Section (iii): Shoemaker and Wiggins: a preference and a qualification.

We tried in the last section to make plausible the idea of bodily continuity being a necessary condition (and hence a "privileged" criterion, see pp 110 - 111 above) of personal identity. But, it was argued, although Williams would be substantially right in holding this he would be wrong to insist that bodily continuity must be understood only in the sense of the continuity of the whole body. It is because of the latter, it was pointed out, that Williams failed to explain consistently a puzzle situation - or rather a type of puzzle situations ( e.g. whole brain transfer cases) - which he nonetheless felt obliged to regard as plausible and reasonable. With a view to making Williams' account consistent, it was suggested that the spatio-temporal continuity of any part of the body should be counted as bodily continuity, since what matters in bodily continuity is spatio-temporal continuity which here would be well preserved. But even by Williams' own admissions bodily continuity (even in his sense) is not a sufficient condition of personal identity, "and other considerations, of personal characteristics and , above all, memory, must be invoked".<sup>1</sup> It is for this reason we suggested that - at least if certain plausible and reasonable situations so demand- the continuity of that part of the body, the brain, must explain the identity of persons, which is causally responsible for the memories and other mental capacities (we have proposed, for the sake of brevity, to use "personal faculties" to signify all these capacities of persons). As we saw, such modification in the concept of bodily continuity would not only be in the spirit of Williams' main thesis, it would make a more extensive theory of personal identity and would explain not only all normal cases but also some abnormal cases.

Now there may be two prima facie objections to our suggested account

of personal identity which must be got out of our way before we proceed further. The first is that in suggesting this type of explanation we have deviated from the bodily continuity criterion; and the second - which is basically related to and issues from the first - may be that we have espoused a form of conventionalism in proposing (as we may seem to be) different types of explanations of personal identity in different situations.

We can approach the first difficulty by reiterating our earlier claim that in taking brain continuity as our criterion of personal identity we do not essentially diverge from the bodily continuity criterion since the former retains the spirit of the latter by preserving spatio-temporal continuity which, as was shown, Williams lays weight upon for personal identity. It is significant to note that what brain continuity preserves is not spatio-temporal continuity as such (if that can mean anything), but spatio-temporal continuity of something bodily and, in the context of personal identity, something importantly or essentially so. In view of the last, it would be unreasonable to insist that bodily continuity must mean the continuity only of the whole body. For what would Williams say if I lose one or some parts of my body e.g. a hand and/or a leg? That the post-bodily-change-person is a different person from the pre-bodily-change-person? I think not. And if I am right I will be also right in saying that what is important is not any actual part of my body but an essential part (or parts) of my body - essential in the role of the bearer of my personal identity. And for well-known scientific reasons, brain is this part of a person's body. Thus, so far as the possession, if not the exercise, of the "personal faculties" is concerned it is the essential part(s) of the body - not the whole body - that matters.<sup>2</sup> That the brain is this essential part (in the relevant sense suggested above) of a person's body is of course a contingent fact; but what is not

contingent is the fact that if it is the essential part then it is a necessary condition of personal identity. Any correct bodily criterion of personal identity should therefore be not only compatible with, but must necessarily include the brain continuity criterion. This is not to say that brain continuity is the same as the continuity of the body (whole body, that is) since, as Williams rightly pointed out,<sup>3</sup> it is absurd to say that the body which now contains Smith's brain is the same as the body which earlier contained Smith's brain. What is claimed is rather that, as a criterion of personal identity, it is brain continuity which does precisely the job that the continuity of the body was supposed to do, (continuity of the whole body minus the causal seat of "personal faculties" would certainly be extremely uninteresting, if not useless, as an evidence of personal identity) and that indeed brain continuity is a form of bodily continuity in so far as it is the spatio-temporal continuity of something bodily.

The position we have been outlining so far and have been claiming to be in the line of Williams, may be described fairly by saying that so far as the situations continue to be normal bodily continuity (i.e. the continuity of the whole body) may be taken to be a necessary condition of personal identity, but that in certain abnormal (but plausible) situations, i.e. if for example brain transferences are successful and are successfully accompanied by the transference of the "personal faculties" the continuity of the brain (which is still bodily continuity) may be taken to be a necessary condition of personal identity. It may be said here - and this brings us to the second difficulty mentioned above - that we are virtually making the criteria of personal identity a matter only of convention. For we seem to be proposing at least two different criteria to apply in two different circumstances, and in doing so we seem to be arguing, like Wittgenstein and Shorter, that if the facts were different, different "geometries"<sup>4</sup> must be called for. Perhaps, there is nothing essentially

wrong with conventionalism - though it is none of my task here to show this; but what I shall try to show is that the position I have been outlining does not have to be committed to conventionalism of any sort. In the first place, we are not proposing different criteria in the different situations envisaged above; for as our arguments in the foregoing and also in the last section would have shown, what look(s) like two criteria can at best be described as only two forms of the same criterion namely, the bodily continuity criterion. Well known examples from recent literature may help to substantiate the point. Material objects like the 'Ship of Theseus' may be reidentified by virtue of the spatio-temporal continuity of their form (i.e. because at every moment of its repair there is a ship of more or less the same form and containing more or less the same materials as the ship of the preceeding as well as of the succeeding moment);<sup>5</sup> or, some material objects like a watch (that has been reassembled after having been disassembled) may be reidentified by virtue of the spatio-temporal continuity of their parts.<sup>6</sup> But this is not to say that completely different criteria for the identity of material objects are being offered in different situations; the two criteria are only two forms of the same criterion, i.e. spatio-temporal continuity.\* Moreover, we do not, in our brain-transfer case, say anything to the effect that this is what 'being the same person' should be called in this situation. If we did say that we would be meaning something essentially different by the phrase "being the same person as" in this case. But I do not think this would be the case in view particularly of the fact that our criterion of 'being the same person' is essentially the same in both

---

\*It will do well to recall in this connection that, for similar reasons, we suggested in Ch 2 sec (ii) that Butler's famous distinction between the two "senses" of identity should be more accurately read as the two kinds of "being the same".

the cases, namely, the continuity of that part of the person's body which is causally responsible for what we have called "the personal faculties".

If what has been said in the foregoing is correct, it follows that a consistent theory of personal identity along the lines of bodily continuity can be maintained to account for personal identity both in normal situations as well as in some abnormal situations, like Shoemaker's brain-transfer situation. We have seen, of course, that Shoemaker intended this crucial example to invite our attentions precisely to the opposite of what our analysis of the example has pointed to. For firstly he believed - not with any good reason, as we saw - that it is "absurd" to say that brain-identity is the criterion of personal identity, and secondly he believed that - again mistakenly, if our analysis is correct - this example, plausible as it is, obliged us to diverge from the bodily identity criterion, and to stress, instead, on the memory-and-character continuity criterion. (That Shoemaker's example points to the importance of the memory-criterion is fairly apparent, but that it proves this is anything but clear. On the contrary, our analysis has shown that if, in such cases, we are obliged to follow the memory-and -character-traits, it is because of the continuity of a bodily part - the brain - on which the former is thus dependent.) One reason for drawing the last conclusion from the case in point may be the popular preoccupation with the idea that bodily continuity is the continuity of the whole body (and clearly this does not hold in the present case); but I do not think that Shoemaker would be preoccupied with this idea at least he would not be consistent if he held it as a necessary fact. For there are indications in what he in fact has said that brain continuity can be a form of bodily continuity. E.g. he remarked,<sup>7</sup> if we said (in the case in question) that Brownson is Brown then "we certainly would not be using bodily identity as our criterion of personal identity. To be sure, we are supposing Brownson to have PART (Shoemaker's emphasis here) of Brown's body namely, his brain",

though he ruled out as absurd the suggestion that brain identity might be our criterion of personal identity. The remark under emphasis above has all the indications that Shoemaker would not object to brain identity being a form of bodily identity. If this is so, what could be his reasons for objecting - as he expressedly did - to its being a criterion of personal identity? One possible reason may be - though I shall be very surprised if it were his reason - that even if brain is a part of a person's body yet the continuity of this part is not very relevant to personal identity. For if the continuity of this part of the body is not relevant to personal identity, that of which part is? And to this may be added the result of our earlier argument in this section that the continuity of all the parts of the body is not necessary, nor even relevant, to the identity of a person. Another reason, which Shoemaker in fact gives, is that if we accept brain continuity as our criterion of personal identity then it would have the absurd consequence of insisting that Brownson must be Brown even if the result of the thought-experiment were just the reverse and Brownson were to have not only the body (minus his brain) of Robinson but also the latter's memories and character etc. We argued in the last section that this result would perhaps be absurd, but that it would not be a result we are bound to accept; we might, instead, more reasonably say that the new person was Robinson which, it was shown, would have pointed to the necessity of bodily identity anyway. What is important is the continuity of the causal seat of the "personal faculties" which, in the counter-counter-factual (if we may so put) situation, would not be the brain but some other part (we may not know which) of the body. Sure, we would in that case be parting company with brain continuity, but we would not part company with the principle behind it - the principle that lies behind our interpretation of the bodily continuity criterion of personal identity.

Now, whatever else might be said to have been Shoemaker's reason for inventing that famous case<sup>and</sup> for drawing from this the conclusion he did draw, it will not be said that it was any preoccupation with the "memory theory" that prompted him to invent the case, the likes of which have customarily been used by 'memory theorists' to serve their purpose. As we shall see, although he was inclined to draw from it a conclusion that suited the memory theories he certainly did not want to generalise this (see especially Self-Knowledge and Self-Identity p 194); on the contrary, he has pleaded for the primacy of the bodily criterion and given persuasive arguments to show that while the similarity of memory-and character traits etc is a criterion of personal identity, it is nonetheless dependent on the bodily criterion for its use. As a matter of fact, one of the main features of his account of personal identity is to argue against the supposition that the criteria of personal identity are 'non-physical'\* or mental. We shall presently see that his position last mentioned is generally sound; his approach, though, to his brain-transfer case and to the alleged cases of bodily-transfer in general - namely, his unhesitant concession to a purely non-physical account of personal identity in these cases - remains a characteristic enigma. More will be said later on about such cases and Shoemaker's approach to them. But now to an assessment of his more general position.

Shoemaker attacked the view that the criteria of personal identity are "non-physical" by attacking two (more) general suppositions which might have contributed to the plausibility of that view. The first is the supposition (issuing from Descartes) that the relation between mind and body, and hence between the mental and the physical features of persons,

---

\*By saying that the criteria of personal identity are "non-physical" is usually meant by the proponents of that position, and I shall so use it, to mean that the <sup>memory</sup> ~~my~~ criterion is the sole criterion to the absolute exclusion of bodily identity.

is purely contingent. The second supposition, due partly\* to the writings of Reid and Butler, is that the real criteria of personal identity are the criteria one uses in making identity judgments about oneself or in making statements that imply one's own identity through time: since in making such judgments we apparently make no use - or at any rate don't have to make any use - of facts concerning our bodies, this supposition thus lends support to the view that the criteria of personal identity are "non-physical". Almost the entire spirit of Shoemaker's thesis has been to try and show that this supposition is mistaken. But it has also been an important part of his thesis to fight the first supposition that the relation between mind and body is purely contingent. We shall be presently concerned with his treatment of this first supposition and take up the second thereafter.

It seems an obvious enough fact that the possession of mind is what constitutes an essential feature of persons as distinct from other things.<sup>1</sup> And if this is so, the (first) supposition that mental and bodily states of persons are only contingently related might lend support to the contention that the criteria of being the same person are non-physical or mental - or at least essentially so. (That they must be solely mental or 'non-physical' follows from the second supposition outlined above.)

The thesis that the relation between mental and physical states of persons is contingent has at least two versions. The first is the weaker

---

\*"PARTLY", because although Reid and Butler held that the real nature of personal identity can be known from one's own case, they were emphatic in denying that there are any criteria of personal identity (See ch 2 sec 11 above).

<sup>1</sup>I am not suggesting - much less argue - that non-human animals cannot be said to have minds, though I have given reasons earlier (Ch 1 sec 1) for supposing persons to be still distinct from animals by virtue of their having what were called "higher-order intuitions". However, the possession of minds by other animals is not incompatible with this feature being essential to persons; and in view of the said distinction between persons and animals the possession of minds by the former would be, so to say, importantly essential.

version (let us call it the weak contingency thesis - or the WCT) that any particular mental state is logically independent of any particular bodily state, that the former could exist independently of the particular bodily state with which it happens to be correlated. But there is a much stronger version (call it the strong contingency thesis - or the SCT) to the effect that although mental states happen to be correlated with physical states it is conceivable that any mental state could exist in absolute independence of, or without at all being correlated with, any bodily state whatsoever.

Understandably, Shoemaker addressed himself mainly to the WCT, for if one succeeds in denying this that would (automatically) result in denying the SCT. i.e. if the relation of a particular mental state (seeing, for example) to the particular bodily state (the condition of the eye, or simply, the eye) which normally is required for the former to be effected, is shown not to be contingent, that would make its relation to that bodily state, and hence by implication to the body, logically necessary; and the same will go for all mental states. As we consider his arguments it will appear that while Shoemaker's intention was to deny the WCT the net outcome of his arguments was not so much its denial as simply the demonstration of the (rather obvious) fact that even if it <sup>is</sup> not denied the SCT will not follow.<sup>8</sup> And this being so, it will be said, not only did Shoemaker's arguments hardly serve any purpose, but the entire argument in this connection was to no purpose.

The WCT appears to be very plausible from the following considerations. To take Shoemaker's crucial example: usually it is almost an analytic truth that people see with their eyes. But it may be imagined that a person, instead of seeing things from the front of his face (i.e. with his eyes), could see them from behind him - that he could do this in a perfectly normal way (i.e. without using a mirror or any other artificial device) and could sincerely say, "I see a tree in front of me" while in fact the

tree is (what normally would be called) behind him. And if this happens, it would seem to follow that the relation of seeing (something mental) to the point of view, the eye, from which one sees (something bodily) is merely contingent. But if this happens, Shoemaker rightly points out, then, if it is not to be a case of clairvoyance or any other like abnormal ability to make true perceptual statements, it should always be possible to discover some point on his body wherefrom he can be said to be seeing and which would be the "point of view" from which he can be said to be seeing. A "point of view" there has to be in order for his perceptual statements to be true (or for that matter, to be conceivably false). And this point of view will have to be on the seer's (perceiver's) body so that what Shoemaker calls his "ego-centric statements" may have any sense at all. Such statements, which indeed make our perceptual (visual) statements precise and informative in the way they are, are the statements like "I see a tree in front of me", "I see a chair on my left" etc. If the crucial "ego-centric" words in these statements, like "in front of", "on the right", etc., have any significance at all it is because of, or in relation to, the position of the seer's body and more precisely of the point of view on his body. Thus, if seeing requires a point of view and that too on the seer's body, it follows that the relation between seeing and the seer's body cannot be contingent. This line of argument has not shown, however, and I don't think it can - that the relation between seeing and the point of view from which things are seen are not contingent, i.e. that the WCT is false. But one need not show that the WCT is false, in order to show that the SCT is false, for clearly the truth of the WCT is still compatible with the falsity of the SCT. Nevertheless, for the reasons given earlier, Shoemaker wanted to show the falsity of the WCT. More particularly, he was anxious to show that it is logically impossible for the point of view to be constantly shifting its position on the seer's body. Clearly, if he had succeeded in showing this he would have succeeded

in denying the WCT and, with it, the SCT. But we shall see that his argument for this is far from convincing. And moreover, the apparent motive with which he tried to attack the possibility of such "shifting" indicates a fundamental confusion and consequently makes it unclear as to what it is he was trying to prove (or prove false). For he is found persistently to be claiming that if the relation between seeing and the body were contingent then "it ought to be possible for the point of view from which a person sees to be constantly shifting its position on his body"<sup>9</sup> (my emphasis). But one fails to see why the latter ought to be possible - how in particular, this is a consequence - much less an important consequence - of the SCT (i.e. of the theory that the relation between seeing and the body is contingent); since ex hypothesi the point of view is to be shifting on the body. What it is a consequence of is the WCT, and if Shoemaker thinks, as most palpably he does, that it is consequence, even an important consequence, of the SCT, then he is certainly confusing the two theses of contingency which I have shown to be clearly distinct from each other. If it is plausible, which certainly it seems to be, that the point of view may be thought to shift its position constantly on the seer's body, clearly no reasonable Cartesian would argue from this that the relation of seeing to the seer's body is purely contingent; and if anyone did, he would be making an inadequate inference, and Shoemaker knows that; but in making the above claim expressedly and persistently he has left his readers wondering if he knows it well enough. For if he did, he would not be spending so much of time and space in fighting the WCT. To be sure his success in refuting the latter would have paid him well (in refuting SCT). But we shall presently see that he did not quite succeed in this attempt.

The burden of his argument has been that if the point of view were allowed to shift its position constantly on the seer's body - now on his face, a moment later, at the back of his skull, then on his side, and so

on - then one would not be able to express the location or position of the objects one sees by means of the "ego-centric" statements. But here again, two types of constant shifting should be distinguished. First, it may mean an infinitely rapid (Heraclitean) change of the point of view so that it may not in principle be possible to check upon the truth of an ego-centric statement like "I see a tree in front of me" since in that case, what is my "front" now is no longer so in the very next moment (or fraction of the moment). To suppose that the point of view be constantly shifting in this way is certainly "to make the cash value of the statements, 'there is a chair directly in front of me', 'I see a chair on my right', and so on the same as that of the pure existential statement 'there is at least one chair (somewhere in the universe);'"<sup>10</sup> and to do this would be to obliterate the distinction between ego-centric statements and other kinds of statements, and the distinction between perceptual knowledge and clairvoyance.<sup>10</sup> But it is difficult to see that this has to be the inevitable consequence of the WCT. For there could be, I suppose, a second type of "constant" shifting of the point of view which, if true, would still establish the WCT. And this is the type (or rather the sense) of "constant shifting" which indeed Shoemaker himself prescribed (see Ibid pp 179 and 182), namely that there should be certain regularity in the change - that between every two changes of position of the point of view, there should be a period long enough for the subject to be able to make a number of ego-centric statements all of which would be true/false with respect to a particular position of the point of view. Once this is allowed the WCT would still hold, and Shoemaker's long drawn arguments against it would be seen to be not only needless but also fruitless. Moreover, Shoemaker certainly can't specify (or legislate?) how long this "long enough" period must be and how "less constant"(?) the constant shifting should be in order for his lawlike prescription may be effected.

It is fairly conceivable that a man could see from different "points of view" at different moments, and that his "ego-centric" statements be conceivably tested with respect to the particular moment of perception. The fact that the "ego-centric" statement "I see a tree in front of me", made at moment  $\underline{m}_1$  with respect to the point of view  $\underline{p}_1$  can't be true at moment  $\underline{m}_2$  with respect to the point of view  $\underline{p}_2$  need not be disturbing - need not present a logical difficulty. For after all, the statement "There is a tree in front of me" made at time  $\underline{t}_1$  is not true when I look aside at time  $\underline{t}_2$ , yet that does not rule out my looking aside. As Strawson rightly pointed out "could not vision swivel as a lighthouse beam, not the light-house, does?"<sup>11</sup> One is not convinced when Shoemaker "explicitly asserts that while one could give sense to the notion of someone's seeing from the back of his head, it could make no sense to suggest that the point (in one's body) from which one sees could be constantly shifting" (Strawson op cit, his emphasis). Moreover, as Strawson also pointed out, it is "unduly restrictive" when Shoemaker rules out the possibility of all round vision while giving "no argument to show why this is any the less of a possibility of all round hearing".

Thus, it would seem, Shoemaker's argument against the contingency of the relation between mental and physical phenomena is unsure and rather shaky and accordingly, his argument against the contingency of relation between personal identity and bodily identity needs stronger support than what he offers. The "shaky" nature of his arguments can be clearly seen from what he says in pp 193-4 (Ibid):

"Earlier in this chapter I said that if certain things occurred, we might say, without absurdity, that a person is able to see from some part of his body other than his eyes. But we saw that to admit this is not to admit that there is only a contingent relationship between seeing and the body, since if the latter were the case, it ought to be possible for the 'point of view' to be constantly shifting its position on his body. Likewise, to say that under certain circumstances we might say that someone had changed bodies is not to say that there is only a contingent relationship between personal identity and bodily identity, since

if the latter were the case it should make sense to suppose that someone might be constantly changing bodies. ... It can be shown, I think, that this does not make sense".

But one is not convinced as to why it does not make sense. As we have seen, if it makes sense to say that a person can see from the back of his head, it also should make sense for the 'point of view' to be constantly shifting its position on his body. And likewise, if it makes sense (as, presumably in Shoemaker's brain-transfer case) for a person to inhabit different bodies at different times, it should make sense to suppose that someone might be constantly changing bodies. Perhaps the impossibility envisaged by Shoemaker is practical or factual i.e. that it just does not happen or even that it could not be known to have happened. But if it is not logically impossible for points of view to shift constantly then this is sufficient to weaken Shoemaker's argument together with the thesis it purports to establish. If facts and principle conflict, so much the worse for the facts.

To add to the "shakyness" of Shoemaker's arguments, notice further that the analogy given in his passage above is not quite fitting to his purpose. For although the supposition of someone's changing bodies (constantly or otherwise) will show that the relation between personal identity and bodily identity is purely contingent, the (analogous) supposition of one's 'point of view' constantly changing its position on one's body will not show that the relation between seeing and the body is only contingent. Yet the latter seems to be the paradigm of Shoemaker's demonstration of the contingency between seeing and the body.

However, **my** critical gesture towards Shoemaker's arguments against the contingency hypothesis and the consequent "non-physical" account of personal identity should not be taken to mean any commitment, on our part, to Cartesianism or to the related view that the criteria of personal identity are "non-physical". On the contrary, we are inclined to believe as Shoemaker is, that the latter theory is mistaken. Our only intention

has been to point out what, in Shoemaker's arguments, seemed to be needless or even inadequate to show this. This being done, I shall now try to suggest some measures to improve a bit upon his arguments so that they may be seen more easily to prove his point. There are two strands in Shoemaker's arguments. The first is that, although in certain circumstances we might say that a person sees from some part of his body other than his eyes, that does not show that the relation between seeing and the body is contingent. This is true, though what still needs to be explicitly shown is that the point of view from which one sees must be situated in one's body. Following Shoemaker's lead, we suggested that this is necessary in order that the "ego-centric" words, which play a crucial role in the making and understanding of our perceptual statements, may have any determinate meaning at all - although we have argued that this did not do the work which Shoemaker wanted it to do, namely to work against the WCT. Now, the second strand in Shoemaker's argument is that while in certain circumstances we might say that someone has changed bodies, that does not show that the relation between bodily identity and personal identity is only contingent, and that a person may be supposed to be constantly changing bodies. We have argued above that this is false, and that if it does make sense to suppose that the former is true then it should make sense to suppose that the latter is true as well. But, it is worth asking, what are the circumstances in which we might say that someones had changed bodies? Obviously for Shoemaker it was the whole-brain transfer cases - the ones we described as exemplifying what we called the "scientific thesis" of change-of-body (see last section). But about such cases it has been argued that "change-of-body" will not be the strictly correct description and, more particularly, that such cases will not show that the criteria of personal identity were solely non-physical. This will therefore not show that the relation between personal identity and bodily identity was purely

contingent - even if such case became of frequent occurrence. If this is correct, Shoemaker, once again, had no need (and in a way this should be unworrying) to fight the possibility of such cases in order to avert its supposed consequence that the relation between personal identity and bodily identity is purely contingent. But he felt the need to reason in the way he did simply because he was not willing to entertain the hypothesis that brain continuity could be our criterion of personal identity. We have seen, however, that his reluctance had no good reason.

There is one more disturbing thing about the alleged cases of bodily transfer which we must pass over quickly. As a possible case, where we might say that a change of body has taken place, Shoemaker would perhaps take a Lockean-Quintonian type of case ("the naive thesis", as we called it) as well, where a complete interchange of the psychological features between two persons may be displayed without there being any explanation of how it had happened.<sup>12</sup> We have said earlier that this type of case can't be plausible cases of personal identity - since they provide no means of <sup>are</sup> telling clairvoyance from memory-knowledge, and/or open to the difficulty of reduplication. And if Shoemaker would still believe and argue that such cases are plausible and do provide any good reason for saying that a change of body has occurred his plea, if any, for this would be very weak indeed. As in his 'brain transfer' case so in this case he cannot prevent someone's constantly changing bodies once he allowed a single case of this sort to be possible. The supposed excuse that these cases would be rare and would constitute exceptions rather than the rule<sup>13</sup> would be only a "lame" excuse and make his case weaker still. (What is the rule anyway? If it is the principle expressed in the proposition that personal identity and bodily continuity go hand in hand, it is certainly disproved, is it not, by, 'in this case (or in some cases) they don't'?)

But let us ask, are these cases possible? That is, are cases of this type possible cases of personal identity?\* At first sight they certainly seem to be, but on closer examination they will turn out not to be. A simple but obvious rule will help us to see this: if a proposition p is logically possible then p cannot lead to or embrace consequences that are incoherent or logically awkward. Now, a "naive theory" like that of Williams' Charles/Guy Fawkes may seem to be quite possible; but if we take it as a genuine case of identity then the consequence is unavoidable that Guy Fawkes is identical with Charles as well as Robert (who might quite possibly have exactly the same identity claim as Charles has) at the same time and also that two different persons (Charles and Robert), in being identical with one and the same person, must be identical with each other, and this, for well known reasons, is a consequence that is logically incoherent. I think, therefore, that a case that allows the logical possibility of reduplication cannot be a plausible case of personal identity and hence must be abandoned. Perhaps it will be said that on criterio-logical grounds these type of cases will point to personal identity - particularly if there are no competitors but only one claimant. But our account of criteria together with our analysis of the memory criterion - which is here at stake - will make us see that this argument fails to do the trick. For in the first place, a criterion is a good reason or justification and, in the absence of any alternative criterion/criteria pointing to the contrary, it will serve as our justification (see p101 above). Secondly, we have said earlier (and this point will be further strengthened later) that the memory criterion is used as a

---

\*I don't say that it is impossible for Charles to claim (in the described way) to be Guy Fawkes but that the suggestion or claim that he is Guy Fawkes is certainly impossible, as I argue in what follows in the text.

criterion of personal identity only because it is taken as evidence of bodily continuity\*.

Now, in view of the last, the psychological continuity in the present case will not be a good reason for thinking Charles to be identical with Guy Fawkes - even if he alone has the described similarity with the latter. And in so far as bodily continuity is known to be absent in the present case our first requirement above will leave us with no justification (criterion) whatsoever to make the identity judgment in question. A Butler-Swinburne type of theory may of course claim that even though we have no evidence here, Charles may still be the same person as Guy Fawkes. But we are given no reason whatsoever to say that rather than saying that it is a mysterious case of clairvoyance (See Mackie again, Ibid p 195). As our problem is one of criteria, I conclude with no hesitation that the "naive theses" of bodily transfer can be no plausible cases of personal identity and hence can constitute no evidence for the supposition that the relation between personal identity and bodily identity is merely contingent or that the criteria of personal identity is "non-physical".

Now, another source of the idea that the criteria of personal identity are "non physical" is the supposition, as we noted, that the real criteria of personal identity are the ones we use, not in identifying others but in identifying ourselves (or in making statements in the first person that imply one's own identity). On this supposition it is common enough to feel that there is an essential difference between persons and other material things and accordingly between personal identity and the identity of these other things. But the difference can hardly be brought out by considering the way we know the identity of persons other than ourselves,

---

\*See also Mackie, Problems from Locke, p 190: "Memory and character are not as evidence alternatives to bodily continuity; they serve, in law and in ordinary cases, as evidence of personal identity only in so far as they are evidence of bodily continuity".

since the way the latter is known is hardly any different from the way the identity of other material things is known. Reid has been noted to have said that our knowledge of the identity of other material things and of the identity of persons other than ourselves are based much upon the same grounds. (The naturalness of this supposition has been discussed at length in Ch 1 (ii) and also in Ch 2 (ii), and so will be only assumed here in order to avoid repetition). The real nature of personal identity and the criteria, if any, for it can therefore be known, it is argued, by considering the way identity judgments are made and known in one's own case. But when one considers how one comes to know one's own identity through time one does not - and does not have to - consider any facts about one's body.\* For if one can be said to know one's own identity through time, this is, or can be, done by the fact of one's remembering that one did certain things in the past or by the fact of one's being able to make and know memory-statements in the first person, and it seems hardly necessary that one should know things about one's body in order to be able to make and know such statements. Thus, if the real nature of personal identity/<sup>is known in self-identity</sup> (Reid and Butler) and if the latter is known or can be known without using any bodily criteria, the conclusion might naturally be drawn that the real criteria of personal identity are non-physical. This sort of conclusion seems all the more natural in view of the fact, which seems necessary, that personal identity (the paradigm case of which is thought to be self identity) is known in (personal) memory and is reflected in the first person memory-statements.

Shoemaker's major concern<sup>14</sup> has been to try and fight this conclusion and this he does by showing that the conclusion does not follow from its

---

\*Once again I will be very brief on this point since it has been discussed at length before.

supposed premiss. More particularly, he has argued that while it may be true that one does not use bodily identity as one's criterion of personal identity when one says on the basis of one's memory that one did such and such in the past, this does not mean that one uses something else (some non physical fact) as the criterion; the fact rather is that, he said, one uses no criteria at all in such cases.<sup>15</sup> This explanation of the situation seems fairly convincing and we have already said<sup>16</sup> that if, in making first person memory statements, one does not and need not use bodily identity as the criterion of personal identity, one cannot be said to be using any other criterion either. For what the "other" criterion could be but memory? And if so, it was asked, do I have to remember that I remember that I did such and such things in the past? But does not this breed the further problem of how this second memory has to be evidenced and so on ad infinitum? Most prominently, Shoemaker has argued that if my memory-statements that I did such and such thing were based on a criterion of my identity, then it should be possible to discover, by the use of that criterion, that the person who did the action was not myself. But clearly, this is not possible. For if I remember doing a thing then I remember my doing the thing,<sup>17</sup> and nothing could show - except my remembering being misremembering - that the remembered act was not done by myself or that the doer of that act was not myself. My remembering could, of course, be a case of misremembering, but to show that it is, one would require the evidence of my bodily facts which ex hypothesi is not in use here. In other words, in the type of cases we are concerned with, we do not use the bodily criteria and cannot use anything else (viz mental criteria) as our criteria of identity in order to know the truth of our statements and consequently to know our own identity which these statements reflect. It follows therefore that the second supposition, even if true, does not inescapably lead to the conclusion that the criteria of personal identity are "non physical".

After having thus shown that the two popular suppositions (as outlined above), which seem to lead to the conclusion that the criteria of personal identity are "non physical", are either mistaken or misunderstood, Shoemaker goes on to argue that the supposed psychological criteria though they are criterial, can't be the sole - not even the primary criteria of personal identity. Two sets of such criteria have been suggested: (1) similarity of character, personality, interests etc, and (2) similarity of memory claims.

About the first, Shoemaker says that it is neither logically necessary nor logically sufficient condition of personal identity. It is not necessary, for the character, interests and personality etc, do change in a person and yet we do not call him a different person for that matter. It can't be sufficient since it is possible for two persons (or more) to have exactly similar character and personality etc (e.g. Williams' example of Charles-Robert/Guy Fawkes). Further, from the fact that a person's character and personality do change it follows that there are other criteria of personal identity which tell us that they do. And these other criteria must obviously be the bodily continuity. (At this point one might notice an oversight - if not an error - in Shoemaker's arguments, sound though they are. For the other criteria by reference to which we may judge that a person's personality etc have changed might well be the memory-claims that he makes, instead of his bodily continuity being the 'obvious' candidate. The answer to this is that while one could use a person's memory claims as ground for saying that his character and personality etc have changed, yet this would eventually point to the same conclusion, for Shoemaker has strong arguments in the immediately following section (Ibid Ch 5 sec 10) to show that the use of the memory criterion inevitably points to and depends on the use of the bodily criterion. This will be brought out in the following paragraph). Furthermore, in order to find out anything about a person's character, personality etc - whether

they have changed or not - we have to observe him over a period of time and what we observe is how he acts, talks and behaves, <sup>this means that</sup> and/we have to observe One and the same person. But how do we do this? Not by observing the similarity (or otherwise) of his character, personality etc, for that is what we are trying to find out about. Thus here "observing the same person" could mean nothing other than "observing the same body". It follows therefore that the use of similarity of character and personality etc as a criterion of personal identity at all presupposes the use of the bodily identity criterion.

As regards memory, Shoemaker argued that it fares no better and that its use as a criterion of personal identity would similarly depend on the bodily identity criterion at some stage or other. First, there is the question of determining whether a person understands, and is able to make correct use of the word "remember" and its cognate words. It would of course be an utterly boring enterprise if we start questioning the genuineness (in this respect) of every or even most memory claims that are made by people, and very few, if any, actually do this; it is rather natural that people's memory claims (or what sound like ones) are taken as memory claims at their face value. But when it comes to the question of determining whether the memory criterion is the sole criterion or even a criterion that is logically independent of the bodily criterion of personal identity, the question does have its logical significance and needs to be settled. It is true of the word "remember" (as also of any linguistic expression) that in order to establish whether a person understands it and its cognate words and uses them correctly, we need to observe his linguistic behaviour (involving these words) over a period of time. And this requires observing his bodily and behavioural facts over that period, and as such presupposes the bodily identity criterion. In the second place there is the question of checking the genuineness of someone's memories as opposed to mere memory claims. For a person's memories may be mistaken - however sincere

they might be. (Once again checking upon people's memory claims is not a usual phenomenon, not so even in most cases: Shoemaker even went on to show, it is a necessary truth rather than a contingent or an inductive one that people's sincere memory statements are generally true; see Ibid, Ch 6 secs 5 and 6 . But the question of how we would check upon the truth of a memory claim is nonetheless important). Yet for the purpose of this checking we can't use the claimant's memory as our criterion of identity. Nor will it do to make any other psychological facts (e.g. similarity of his character, personality etc) as our evidence that this person did what he "remembers" doing, for this, as we have just seen, can be used as a criterion of personal identity only if bodily identity is a criterion, and this will undermine the thesis, which we are concerned with, that memory criterion is the sole criterion of personal identity or is logically independent of the bodily criterion. Moreover, if it is involved in the meaning of the word "remember" that the person remembering doing an action must be the person who did that action, the checking of a memory claim would consist in finding out whether the "rememberer" is the same person as the one who did the remembered act. And the criterion of personal identity which is needed here would have to be the bodily identity criterion, since, as we saw, neither the similarity of memory claims can do (for that is what is suspect) nor that of personality and character will usefully do (since this points to the bodily criterion anyway). It follows therefore that the memory criterion (or the psychological criteria in general) cannot be the sole criterion of personal identity in absolute exclusion to the bodily identity criterion and that, being logically dependent on the bodily criterion, it cannot even be said to be the primary criterion. It is for this reason that Shoemaker would prefer the bodily criterion and regard it as the more fundamental or the primary criterion. Indeed he claimed that "if it (bodily identity, PKM) were not a criterion, nothing else would be evidence of personal identity."<sup>18</sup>

Shoemaker's reasons for saying the memory continuity is nonetheless a criterion is predominantly that the puzzle cases (which we have called the "scientific thesis" as well as in the "Naive theses" of change of body) so much incline us to say that change of body has occurred, or at least leave us in doubt as to what to say,<sup>19</sup> but our reason for saying this would be slightly different from Shoemaker's, for we want to say, for the reasons given before, that if in certain puzzle cases (i.e. only in what we called the "scientific" cases) we have reason to say that a change of body (in the relevant sense we envisaged earlier) has occurred, then memory must certainly be a criterion of personal identity. But saying this, we argued, does not make memory the sole, or even the primary criterion, and as we have shown, the latter's being a criterion can reasonably be said to be due to (see last section) the bodily continuity criterion. If what we have said is right, Shoemaker has not given any good reason for his belief that at least in certain cases (i.e. these puzzle cases) memory continuity could be the sole criterion of personal identity and personal identity judgment in this case is made on purely non-bodily grounds. Moreover, if he is right in arguing that any psychological criterion of personal identity is not logically independent of the bodily identity criterion, then reasoning, as he did, in the puzzle cases would expose a serious inconsistency in his position. As I said earlier his analysis of the puzzle cases remains a characteristic enigma in his account of personal identity.

Now, our discussion in the last section as well as in the present, and our analysis of the puzzle situations which so much captured Shoemaker's imagination and even his reasoning, has made us see the importance of two factors concerning the reidentification of persons. First, in order that memory continuity may serve as a criterion of personal identity, it has to be understood as a causal notion; and secondly the causal chains linking

the memories with the remembered actions or experiences must be spatio-temporally continuous (see last section pp 123-4 above). But for this last condition to be satisfied it seems to be required that the capacity to remember and all other mental capacities (which together we have called the 'personal faculties') be seated in and carried through some parcel or parcels of matter; without the continuance of some such thing, that is spatially locatable and traceable, it seems the spatio-temporal requirement will not be fulfilled. This will bring us to the position recently advocated by Wiggins in his Identity and Spatio-Temporal Continuity. After showing that spatio-temporal continuity is a necessary condition for identity in general, he rightly argued that "spatio-temporal continuity" is incomplete if it is not specified what it is the spatio-temporal continuity of. This requirement can be described as the "covering concept requirement" of identity. Obviously the covering concept here is person. It is thus an inescapable part of Wiggins's thesis that a criterion for the identity of persons must be potentially analytical of the concept person.<sup>20</sup> It is a conceptual truth that being a person distinctively involves not simply having a mind or possessing mental capacities but also, more particularly, being able to remember a considerable amount of one's past. This has been emphasized in most writings on the subject, from Locke to the present day. And it seems particularly true in the present context, because if people did not remember something at least of their past,\* we would not now be talking of such a thing as personal identity. For it is by remembering some of our past deeds and experiences

---

\*I am not ruling out, as Locke's definition of personal identity would, amnesia as logically impossible. My analysis of 'person' in this respect simply amounts to the claim that to be a person is to be able to remember some of one's past for some period of one's life - not that one must be doing so at all moments of one's existence.

that we are aware of our own identity through time, and this also creates in us the presumption that other people, who are so much like ourselves, must be aware of their own identities in the same way.\* This at least is the kernel of the truth that personal identity and memory are logically connected. Thus Wiggins rightly describes these capacities as the characteristic functioning. But if we were to understand these capacities as something purely mental or non-physical, then we would not be able to apply the spatio-temporal continuity criterion to them; yet it is the spatio-temporal continuity of precisely these capacities that we are supposed to determine when we are determining the identity of a person. It follows therefore that if we are to specify the spatio-temporal continuity under the concept person, we cannot take the continuity or similarity of mere memory claims (and/or other mental capacities) unless it had or was supposed to have a physical basis. We must therefore look for some physical (i.e. bodily, since we are considering the identity of persons) part(s), if any, which is responsible for a person's memory and character etc, and the continuance of which would explain the continuity or similarity of the latter. For well known reasons the brain can be chosen for this role, for the brain happens, as a matter of scientific fact, to be causally responsible for the possession, exercise and the continuity of what Wiggins has called the "characteristic functioning" of a person. Thus he said,

"It would be better, after a conceptual analysis of the essential and characteristic vital functions to analyze person in such a way that coincidence under the concept person logically required the continuance in one organized parcel of all that was causally sufficient and causally necessary to the continuance of essential and characteristic functioning, no autonomously sufficient part achieving autonomous and functionally separate existence."<sup>21</sup>

Clearly Wiggins takes as the criterion of personal identity the continuity of the brain, or of that bodily part<sup>22</sup> (or those bodily parts) which

---

\*That the latter is a logical precondition of the former has been argued in Chapter 1 section (ii).

plays the role of the "seat of memory and consciousness". Because of the reasons we have been pursuing in the last section as well as in the present, this is a healthy modification in the bodily continuity criterion and is quite an acceptable account of personal identity. The beauty of Wiggins' account is that it purports to "logically exclude" the possibility of reduplication which has often been thought to vitiate even the brain continuity criterion. For the brain could be bisected and each half be transplanted in a different body with (it is supposed) the consequent transference of all the memories and character etc of the donor to each of the recipients; our calling the claims of each recipient person "memories" will be justified on exactly the same ground (i.e. the causal link) in this case as in the simpler case of whole brain transference. Wiggins tried to rule this out by insisting on the "continuance in one organized parcel of all (my emphases) that is causally sufficient and necessary"; and 'having all that is causally responsible for A's memory and character' is a logically one-one relation by means of which only one person could be related to A. The same purpose can be served by taking the continuity of the whole brain; for 'having the same (whole) brain\* as ... ' would be equally a logically one-one relation. For this reason when earlier we pleaded for the brain transfer cases being plausible cases of personal identity we meant to be, and should be, understood to mean whole brain transfer cases. As regards the supposed split brain transplant cases, we want to suggest that they will not be plausible cases of personal identity, in the sense outlined above (see p 56-57). For not only the supposition of such cases is embarrassed by utterly awkward consequences, it is also

---

\*Indeed, not only saying "having the same whole brain as .." is odd and unnatural, but saying "having the same brain as .." very naturally suggests that we are talking of the whole brain. On the contrary, when the possibility of split brain transference comes to the picture, it would be rather more natural and accurate to talk of "having the same brain part as.." or something like that.

committed to consequences that are logically incoherent. The awkwardness lies in the fact that there would be no "logical limits" - to borrow Williams' chosen phrase<sup>23</sup> - to how far we can go on with such possibilities: should we stop at bisection or go on to speculate the possibilities of trisection or even multisection of the brain? The decision where to stop is bound to be arbitrary, and as Williams has rightly warned us,<sup>24</sup> no account of personal identity can be philosophically satisfactory if it has to depend on an arbitrary provision for its success. The more serious difficulty with such supposed cases is that, worse than the "naive theses" of bodily transfer which we have shown to be implausible as cases of personal identity, these cases necessarily carry with them the possibility ex hypothesi of reduplication which, in its turn leads to the logically incoherent consequence that two different persons (the recipients of each half brain), if they are to be identical with the brain donor, must be one and the same person. Thirdly, no determinate identity question can arise in such cases. For as Wiggins rightly pointed out, the proper form of an identity question is to be in terms of a suitable covering concept (see Ibid esp p 56), and given that x is the brain donor and that y and z the recipients of each half brain, what are we to ask or judge to be identical? X is the same what as Y and Z? He cannot be the same person as y and z, nor can he be said to be the same persons or pair of persons as y and z.<sup>25</sup>

Thus, in the light of the foregoing Wiggins' account of personal identity seems logically sound\* and it has the unique advantage of steering

---

\*Though, unhappily, Wiggins in fact does not especially give a logically sound suggestion to deal with the split brain transplant cases. For he seems to suggest, though not entirely unequivocally (see ISTC pp 55-56), that there will be personal identity in such cases only if one half of the brain takes and the other half destroyed. Besides being more pointedly vulnerable to Williams' warning mentioned above, this type of solution has the absurd implication that who one is depends on what happens to someone else. I believe further that this is a prima facie strong objection against the current theories of so called "non-branching" (psychological) continuity.

clear of the two supposedly conflicting criteria of personal identity which seem prima facie so much incompatible; for on this analysis, a correct memory criterion and correct bodily (spatio-temporal) criterion will necessarily coincide (as he says in p 45 of ISTC). Even Shoemaker, who embraced multiple criteria and believed in there being genuine cases of conflict of criteria, was forced to see the this truth in Wiggins' analysis.<sup>26</sup>

However there is at least one thing that remains to be explained and that does not come out clearly in Wiggins' account; and that is, why is it that the 'seat of memory and consciousness has to be located in the body of the person. That this is always, or normally always, the case and that it is never known to be otherwise is perhaps a good reason (though not conclusive reason) to suppose that this must be the case. But a stronger reason can be given by saying that if it were not the case, our first person psychological statements would not be informative in the way they are. One important kind of these statements are what we have seen to be the "ego-centric" perceptual statements (Shoemaker). We have seen earlier in this section that these statements - the "ego-centric" words, in particular like 'in front of', 'on my left' etc - will not 'have any determinate meaning if the point of view from which we see or perceive were not located in our bodies. Thus it would be necessary or a conceptual requirement for those words and statements to have meaning and meaningful application that the truth of our perceptual statements (made in the first person) must be related to the described fact about our bodies. Now the second important kind of our first person psychological statements are the memory statements that we make in the first person. It can be shown, I think, that the truth of these latter statements must be consequently related to certain facts about our bodies. For it is a logically necessary fact that our memory statements are related in a certain way to our (corresponding) perceptual statements. And if I remember at t<sub>2</sub> that I

saw a tree in front of me at an earlier time  $t_1$  then the statement (call it M) 'I saw a tree in front of me at  $t_1$ ' cannot be true unless the corresponding statement (call it P), 'I see a tree in front of me now' made by me at  $t_1$  is true. But if, as we saw, P's being true depends on certain facts about my body (namely that the point of view from which I see is located in my body), it will follow that M's being true must also depend upon these facts about my body. This, together with the fact that being a person requires being able to make a number of memory statements in the first person, shows that persons have to have bodies and that the causal seat of memory and consciousness (what we have described as the "personal faculties") will have to be situated on their bodies. This conclusion is further substantiated by the transcendental argument, adduced in the last section, in favour of the contention that the causal link involved in the memory must run through the person's body. If this is so, and if the minimum that matters in personal identity is the continuous possession (not exercise) of the "personal faculties" then bodily continuity, in the modified sense we have given to this concept, will turn out to be a necessary condition of personal identity. It will then be an empirical question as to what (what part or parts of the body) makes the continuance of the "personal faculties" possible and Wiggins' solution is an answer to this question (see also Mackie<sup>27</sup>); but this will not make the continuance of the physical basis of "personal faculties" as related to personal identity a contingent matter of fact.

Section(iv): Derek Parfit: memory with a difference(?)

In the last two sections I have tried to make plausible the theory that bodily continuity - or rather the spatio-temporal continuity of something bodily - is a necessary condition of personal identity, and to show that even though similarity of memory and/or other psychological features is also a criterion, it is due to the former kind of continuity. In any case, it was argued, the psychological criteria cannot be of the sole or even the primary consideration. Quite recently, Parfit has advocated an account of personal identity which, if sound, will totally dismiss our latter conclusion (and since it is dependent on our former contention, our entire position will be dismissed thereby). For his account purports to analyze personal identity in terms of these psychological features or of what he technically calls "psychological continuity", and, implicitly as well as explicitly, bodily considerations are ignored as irrelevant.<sup>1</sup> It seems to me, however, that despite the admitted novelty and ingenuity of its approach to the problem, the theory is largely mistaken. In order to show where exactly it fails to do the trick, I shall examine the theory fully.

The theory offers two important suggestions; one on the nature of personal identity and the second on its importance. The second of these suggestions can be seen to follow as a natural consequence of the first. The first suggestion is that personal identity is a matter of degree - that the identity of persons is not, as it is normally supposed to be, all-or-nothing. The second suggestion is that certain important moral and practical questions, which are usually supposed to presuppose personal identity (in 'all-or-nothing' sense), and give the latter the importance it has can be freed of this presupposition, and that once this is done, personal identity will have no importance. I shall first examine the first suggestion and take up the second suggestion at a later stage of

this discussion.

Ordinarily, questions of personal identity are thought to demand all-or-nothing answers. We are inclined to think that, given our normal criteria of personal identity, a later person  $p_2$  is or is not the same person as an earlier person  $p_1$ . This we think because we believe that the nature of personhood is not fully constituted by all the evidence we can have of it. For example, I am not the same as my body since, obviously, my body lasts longer than I do - or at any rate, we do not last the same amount of time. Moreover, if I lose some parts of my body, I do not cease to be the person I used to be nor do I become only a part of the person (if that may mean anything) I used to be. Similarly, for the psychological features that characterise me as a person. Even if I lose a considerable amount of my memories and character etc. I still continue to be the person I used to be - though I may be said to have undergone a change of personality. Thus, the bodily and/or the psychological features that normally characterise a person are not identical with (or analytic of) the concept of a person. Therefore the fact of personhood does not consist in just these facts; it is a further fact. Accordingly, the fact of personal identity is not just a matter of bodily continuity and the continuity or similarity of memory and character etc; it is a further, deeper fact. Because it is a further fact, personal identity is not affected by the obvious changes in the bodily continuity and the continuity of memory and character etc. e.g. amputated limbs, plastic surgery, on the one hand and loss of memory, change of personality on the other. It is not even affected by the discovery or suggestion of such changes as may purport to show that nothing really is continuing (see Ch 1 sec 11 above). And so, asking for all-or-nothing answers to questions of personal identity does not seem to be an unreasonable demand. In this sense, a person  $p_2$  at  $t_2$  and a person  $p_1$  at  $t_1$  are identical with each other if they have more or less the same

body and if they make more or less similar memory claims and display more or less similar character, personality etc. If not they are not identical, but are different persons.

But Parfit introduced a new way of talking according to which what we call two phases of the same person's life may not be the phases of the same person. A person may dissociate one part of his life from another earlier part. This way of thinking gives rise to a certain view about the nature of personal identity and a certain way of talking. He refers to the former as the "Complex View" and to the latter as the "proposed way of talking".<sup>2</sup> Each is supported by the other. On this view, a man may have an attitude of indifference towards a part of his life; while in such a state he does not regard that part of his life with either pride or shame, pleasure or regret. He might have behaved very badly in his remote past. He may now have undergone some serious change of character, manners and attitude towards things. In such a case, he might say, "I admit that I behaved that way, but the 'I' who so behaved seems to me a stranger." Thus, he may have no regrets now for having once so behaved. This man's attitude towards his earlier self is what Parfit calls "non-identification". Thus it is argued that that man's present self can be different from - not identical with his earlier as well as his later (future) selves. The idea suggested is that we might free ourselves from the dominance of self-interest and anxiety about the future and the past phases of our lives by thinking that one's life is not one single unit but is divided into the many lives of successive "selves". This can be possible by a lessening of what is called "psychological connectedness" which is explained by substantial change in style of life, character or by loss of memory. The more remote is the past 'self' the less is the psychological connection or connectedness between your present "self" and that "self", and therefore the more you are justified in depending (pleading?) non-identification. But, as Parfit himself sees,

this type of non-identification cannot find much defence if the part of past is more recent from which you want to dissociate your present self. "If it is only after some slight change that a man does not identify with his earlier self, such a change would not provide much defence. But if it is after a great change then, I think, it would."<sup>3</sup> This anticipates Parfit's theory of degrees of personal identity.

I shall shortly come to consider the theoretical basis of this view of personal identity. But before that, certain observations seem compelling at this stage. The first reaction to such a view would be that, granted the differences or changes in character, memories etc, which can be supported by facts, it is your present self that is claiming to dissociate from your past self! (The "your" certainly has more than a mere rhetoric force. It cannot be said, either, that my emphasis on this word is tendentious. For the alleged attitude of "non-identification" is surely not of the same type as the one you have towards an earlier "self" belonging to the life-history of some other person. What then is this difference - if not the fact that the earlier self in question is yours?) So how can you cease to care about a period in your past? For even if you have changed, your behaviour then is as much a part of your own past! As Penelhum put it, "That part from which we wish to dissociate our (present) selves is as much a part of us as that with which we identify".<sup>4</sup> In reply to this Parfit says that though the claim is, in a sense, true, it is a superficial truth; "it is like the truth that all the parts of a nation's history are as much parts of its history". He says, after Hume, "what is most important in the histories of nations are the continuities of people, culture and political systems. These vary in degrees. So the identity of a nation over time is only in its logic all-or-nothing; in its nature it is a matter of degree". And he makes similar claims about persons; "what is most important in the survival of a person are a number

of psychological relations. Most of these relations hold, over time, to varying degrees. So the identity of a person over time is only in its logic all-or-nothing; in its nature it is a matter of degree.<sup>5</sup> It is noteworthy that Parfit uses "survival" and "identity" interchangeably in this passage; but this seems not to be in keeping with the general scheme of this theory. For it has been one of his crucial claims that the language of identity should be given up in preference to that of survival.<sup>6</sup> But, nonetheless, he had to use these expressions in this way or else the analogy he makes use of would lose its force, or at any rate, would fail to bring out the desired effect. Now, assuming for the moment that the analogy holds and also that the different phases of a person's life (what Parfit calls different 'selves') are as diverse as the different phases of a nation's history, still it is not clear how this helps one to talk of identity of persons. Why should one claim that the "Complex View", together with its proposed way of talking, offers an account of personal identity? <sup>has given us is a view</sup> All that this view/about the nature of the identity (or otherwise) of 'selves' (in Parfit's sense of the term) - not of persons; and if this view of selves is correct, then "the concept of a person must be wider than the concept (wider than this concept) of a self".<sup>7</sup> If, alternatively, as Parfit suggests, we are to substitute "single self" for "person",<sup>8</sup> then of course it can be claimed that a view about the nature of "personal identity" has been given, but only at the risk of making a considerable departure from the ordinary usage. We have in our language enough room for "person" and even also for Parfitian "self" (or "single self"): as they are used they do certainly overlap, but they never coincide, and so cannot be intersubstitutable. For example we do recognise the distinction between someone's boyhood, manhood and old age - each signifying a single phase in his life history (and hence roughly corresponding to Parfit's "single self"); but each phase, though a part of the person's history, is not the

person. The person is tenselessly each of these and, indeed, all of these together while each of the three phases (three "selves") is the person only as applied present tensedly. In Wigginsian terminology, the Parfitian "self" would be related to the person as a phase-sortal to a substance sortal.<sup>9</sup> Any attempt, therefore, to substitute "single self" for "person" is to blur this important distinction unreasonably.

Another reaction that suggests itself is: can we talk of 'degrees of identity'? - even as applied to the selves in Parfit's sense? Admittedly, what Parfit means by this is that there may be greater difference in respect of character etc between two selves x and y than between x and z, if i.e. y is a more remote past self and z a more recent past self compared to x, the present self. But all it would mean is that z is more similar in character etc to x than y is, and surely to say that z is more identical with x than y is, would be odd enough. And if it is suggested that the problem of personal identity is only about similarities (and not about numerical identity) it would be a very uninteresting suggestion indeed.

Now, let us see what gives a theoretical justification to this new way of talking. There is at least one view, which we explained at the beginning, according to which personhood and personal identity are some further facts, independent of the bodily and psychological connections involved in them. Parfit refers to this view as the "Simple View"<sup>10</sup> as opposed to "his "Complex View". As I have indicated, this view is the generally accepted (though not by philosophers in their speculative mood) view which we normally have about the nature of personal identity. Parfit does not take this view seriously. On the contrary, he believes that the fact of personal identity consists in the holding of certain more specific facts like bodily continuity and the continuity of memory and character etc. he refers to the latter as "psychological continuity". Apart from these specific facts, he does not think that there is any

further fact that constitutes the nature of personal identity. But these facts are found to hold, over time, to varying degrees. So, he thinks, personal identity is only a matter of degrees. We have seen how he takes the help of Hume's analogy to show this point. But it is noteworthy that the specific facts in which personal identity is said to consist is said to be only the "psychological continuity" and that bodily continuity is conspicuously ignored (another Humean feature). The psychological connections are so much emphasised that one almost forgets that the so-called non-identification is effected against the background of the same, continuing body. As a matter of fact, Parfit has expressly said that the bodily continuity, which is involved in personal identity, may be set aside as it is "morally irrelevant".<sup>11</sup> Later on, I shall try to show that bodily continuity cannot be as irrelevant - not even morally - as Parfit thinks it is. But at this stage, even if Parfit is right about his account of personal identity, then the "more specific facts" which are involved in personal identity, and which he thinks are all that matter, include not only psychological continuities but also the continuity of the body, and it would be unjustified to make the nature of personal identity solely dependent on (or even determined by) the former to the total exclusion of the latter. It may be said in defence of Parfit that even if bodily continuity is included in the "more specific facts", it would not affect his theory - since bodily features, too, do change in degrees over a period of time. But even so, a theory of degrees of identity, with its novel way of talking, would not give a right description of "same person" which will be in keeping with our normal understanding of the concept. Then what was the need for the Complex View? What was wrong with the Simple View - which is nearer to our normal understanding of persons and personal identity? With his empiricistic bias - and with his

distinctive Humean way of thinking\* - Parfit seems to have reasoned as follows: the only observable evidence that we have of the identity of a person are bodily continuity and similarity of memory and character etc. (psychological continuity); beyond these facts there is nothing that could show that there is a further fact which the identity of the person consists in. So the claim that personal identity is a further fact, over and above these specific facts, is meaningless, and hence the Simple View is unacceptable. But this in effect is verificationism<sup>12</sup> which needs to be supported by arguments before Parfit's reasons can give his Complex View the status of acceptability. But although Parfit thus seems very much to have relied on the verification principle, he has no argument in its favour. On the other hand, it makes quite good sense - it is certainly not incoherent - to suppose that a later person is the same person as an earlier person even though all the evidence that we have point to the contrary. For there is quite a gap between the (observable) evidence that a proposition p holds and the holding of p. For example "Smith is in pain" may be true even though we have no evidence that he is, and conversely, all evidences may have the tendency to show that Smith is in pain when, in fact, he is not. Neither can Smith himself be said to have evidence that he is in pain when he is. For one does not know on evidence that one has pain. One simply has it. (Is it not what Wittgenstein had in mind when he said in the PV: It cannot be said of me (except perhaps as a joke) that I know that I am in pain?). The fact that we have no evidences for - perhaps we have all the evidences against - Smith's being in pain has nothing to show that the claim "Smith is in pain" is meaningless. Thus, the fact that there can be no (observable) evidence to show that personhood and personal identity are some

---

\*Note especially his claim: "He is one of my later selves and I am one of his earlier selves. There is no underlying person who we both are".  
"Personal Identity" PR 71, p 25

further facts does not make the claim of the Simple View meaningless. Rather this claim has a practical advantage in its favour that it is more in keeping with the way we feel about person-hood and personal identity. Further, the analogy of nations fail in the sense that we can refer to people without making any reference to the nation to which they belong, but we cannot refer to the bodily and more particularly, psychological continuity which is all that matters for Parfit, without making any reference to the person whose bodily and psychological features they are. However, if Parfit is proposing to give an account of personal identity solely in terms of the more specific facts, namely the (bodily and ) psychological continuity, he can be accused of making the same confusion which most writers on the subject have been, recently,<sup>13</sup> charged with, namely, the confusion between the question: "what does it mean to say that a person  $P_2$  at time  $t_2$  is the same person as a person  $P_1$  at an earlier time  $t_1$  ?" and "what evidences can we have for saying that  $P_2$  at  $t_2$  is the same person as  $P_1$  at  $t_1$ ?"<sup>13</sup> As I have been trying to say, the meaning of P and the evidence for P are not the same, for if they were, understanding the meaning of P and understanding the evidence for P would also have been the same. But to say this would be absurd. For understanding the evidence that some one is having toothache is understanding how he behaves and what he says e.g. that he holds his jaw, carefully nurses a loose tooth, says 'oh, it hurts' and so on; but that is not what toothache means.<sup>14</sup> Bodily continuity and psychological continuity (Parfit calls them "more specific facts" in which personal identity consists) are our evidence of two non-contemporaneous persons being one and the same person; but just these evidences cannot express the meaning of "same person". So, to say that personal identity is only a matter of bodily continuity and/ or psychological continuity would be to confuse evidence with meaning. (A fuller discussion of this is given in Ch 1 sec 1). It is true that for the understanding of certain concepts

we have to take note of the grounds on which these concepts are applied - and in the case of the mental conduct words (those signifying 'mental processes') the grounds may be all we can have or rely on. But there is certainly more to the meaning of these concepts than just those grounds. This is evident from the fact that a mental conduct word (e.g. "pain") and an expression signifying all the evidences for the relevant mental process are not intersubstitutable, nor can a sentence containing the one be substituted by a sentence containing the other. This significant linguistic fact speaks faithfully about the nature of things. "An inner process stands in need of outward criteria", (Wittgenstein) but the inner process is not just these outward criteria. This was behind our earlier claim (sec 1 above) that a criterial relation is not a relation of entailment. The criteria for  $x$  do not imply  $x$  and, so cannot be confused with it. If this is true, then a theory which asks us to mean by personal identity nothing more than some specific facts that are involved in our understanding of personal identity and in making judgments about it, cannot be the right theory; and in so far as Parfit's theory asks us to do this, it must be mistaken.

One advantage of the Complex View is sometimes said to be that it makes clear how to describe puzzle cases. Our normal theory of personal identity cannot give clear answers in the puzzle cases because it assumes all-or-nothing answers. On occasions our usual criteria for personal identity may fail to give any clear verdict as to whether  $P_2$  is the same person as  $P_1$ . Perhaps, if an answer is insisted upon, we may have to decide upon an answer. And this is a very uncomfortable position to be in. But if we accept Parfit's way of talking, it is sometimes said, such difficulties would not arise. For on that theory, a later person need not be wholly identical with (or different from) an earlier person; instead one can say that they are "exactly the same", "pretty much the same" or "hardly at all the same"<sup>15</sup> Thus, in a puzzle situation the

answer in one of these formulations would be given, depending on the degree of continuity of the relevant features. "The less and less gradual the brain and other continuities etc the more  $p_2$  is different from  $p_1$ ."<sup>15</sup> But it is evident how uncomfortable you would feel if you are to say, as you must on Parfit's theory, that your father is almost 'exactly the same or 'pretty much the same' or, what seems to be much worse, 'hardly at all the same' as the person he(?) was years ago. Further, irritatingly enough, the element of decision has even an uglier face here. For on Parfit's theory, the identity (or difference) of persons would depend on the extent to which there has been a lessening of psychological connections and the latter inescapably would depend on the decision or choice of the speaker\* who claims non-identification. Curiously, Parfit is committed to this. For he expressly says that where lessening of psychological connections is to be done is left to the choice of the speaker.<sup>16</sup> If this is so, it is not clear how his proposed way of talking fares any better in the puzzle cases than our normal way of talking. For if we must allow room for decision here, then it is at least as bad as the 'simple view' which cannot give clear answers in puzzle situations and may lead us to decide upon an answer. Rather the case of the Complex View, in this respect, seems to be at a greater disadvantage: the decision in this case is more likely to be arbitrary than in the other case for in the latter case who decides is not the person himself but others who may have full factual knowledge of what has happened. It will not only be said that there has been a lessening or

---

\*This will have to be the case since, whatever objective grounds (viz behaviour, including verbal behaviour) we may have for saying that a lessening of psychological connection has occurred in a man, this is still compatible with the man trying to deceive - perhaps with a view to avoiding responsibility of a past misdeed.

continuity of psychological connection, but also that why it has been so (and the latter will have the objective support of what has happened). However, one argument in favour of leaving the decision to the speaker himself may be the fact of the so-called 'privileged access' which one is said to have to one's own identity. But whereas this may be a fact it is not an infallible fact. There may arise certain situations in which a person may not know who he is - or may be mistaken about his own identity, and in such situations others will be required to tell him that. Thus, the idea that who one is may depend on what others say is not all that odd or absurd.

Let us now turn to consider the importance of personal identity and Parfit's second suggestion that concerns this. Parfit thinks that the question of personal identity is important because it is presupposed by some important moral and practical questions (questions about such matters as memory, survival and responsibility). But, he argued, these latter questions can be freed of this presupposition; and once this is done, personal identity would lose its importance. Now, to start with, it is genuinely doubtful as to how these two questions can be set apart, for with the question of personal identity losing its importance, questions of memory survival and moral responsibility also lose their importance. It is not clear how can these latter questions plausibly arise independently of the question of personal identity.

It is an undeniable fact that genuine memory claims do presuppose the identity of the person who remembers. Of course "memory" can be interpreted in such a way as not to make this presupposition binding. Parfit, apparently to make the fact of memory free from such presupposition, takes the help of q-memory (memory of an experience or action which need not be one's own) to replace "memory". The suggestion is that it is possible to q-remember (not necessarily remember in the strong

sense) an experience without implying the identity of the rememberer with the owner of the experience remembered. For instance (with reference to Wiggins' operation, see note 12) the two resulting persons, in the split brain transplant case can q-remember experiences which had happened to the original person, and yet obviously they are not identical with the original.\* Fair enough. But the fact is that in the world as it is, the phenomenon of q-memory cannot totally replace that of memory in the strong sense, and that in this world memory in the latter sense is the more important kind of memory which, hence, keeps the importance of personal identity undiminished. The fact that in our world memory in the strong sense is the most important kind of memory can be seen as a necessary (rather than contingent) fact, and it is this fact which makes memory a criterion or even an evidence of personal identity. I have argued this point elsewhere before (esp Ch 2 ii). I claimed that what makes memory a criterion of personal identity is the fact that "memory" in the strong sense means that it does, for it is this fact which creates in us the presumption that any ostensive memory may be real memory and hence that the "rememberer" may be the same person as the person who witnessed (did, experienced etc) what is "remembered". To be sure, this does not rule out - it rather requires it in order at least to give a non-circular account of the memory criterion - the fact of there being apparent memories or the fact of q-memories in Parfit's sense. But what

---

\*For, Parfit nonetheless rightly argued, if they were they would be identical with each other. But they can't be, because despite their close similarity so far with each other in almost every respect, they would hereafter go in different directions, have different memories and experiences and, quite possibly, may fail to recognise each other when they meet after a long lapse of time (PR 1971 p7).

As he also pointed out, certain moral and practical considerations show the same point: If the two fight a duel and one of them dies, will it be murder or suicide? And will the survivor be one person, half a person or one and half a person? (cf op cit p 8 note 8)

it does rule out, I think, is that the latter be the only form of "memories" in the world. For if it were, our understanding of the word "memory" would be radically different from what it is; and if anything, it would point to the fact that a person "remembering" something is not the same person as the one who did that thing. In other words, it would teach us non-identity rather than identity of persons. Any case of "remembering" (q-remembering or otherwise) would then create in us the presumption that the present person may be different from the earlier person. Thus, in such a world, memory of any form would fall outside the criteria of personal identity and would be entirely irrelevant to it. It follows therefore that a world in which there will be only q-memories that are never memories in the strong sense will not be our world - not even one like it.\* (Yet Parfit is most anxious that his proposed way of talking should apply to ourselves and not merely to his imagined beings.)<sup>17</sup> The same argument would apply against the supposition - which Parfit makes but does not argue for<sup>18</sup> - that q-memory could be the basic form of remembering and that what we call real memory would only be q-memory of one's own experiences (op cit p 16). For one thing, even by Parfit's own commitments the concept of q-memory is a derivative one and is dependent on the basic notion of (real) memory. (I refer to the rather dubious third condition in his definition of q-memory: "I am q-remembering an experience if ... (3) my belief is dependent upon this experience in the same way (whatever that is) in which a memory of an experience is dependent upon it", PR 1971, p 15 my emphasis). Secondly, if per impossible, q-memory becomes the basic form of remembering, then by our above arguments this memory would point to anything but the identity of the person who "remembers". Thus, we see that, in his extreme anxiety

---

\*It is not surprising, therefore, that in the Parfitian scheme of things, identity "does not matter".

to free the memory-criterion from the charge of circularity (cf op cit, p16) Parfit makes it cease to be a criterion at all. But if it has to be a criterion the strong sense of "memory" must be its basic sense and hence the question of memory cannot be freed of its usual presupposition about personal identity. Whence it follows that the importance of personal identity cannot be affected - even by the introduction of "q-memory".

As regards the question of survival, Parfit wants to free this question from its usual presupposition of personal identity by proposing to keep the language of survival and to give up the language of identity. The illustration again is Wiggins' operation. "We suggest that I survive as two different people without implying that I am these people". (op cit, p 8). The two resulting people, by virtue of having strong psychological connections with the original person, can refer to him as "my past self" can claim themselves to be "his future selves". This way of talking is extended to imagine a world where fission is the rule rather than exception, so that in such a world the "descendant selves" can be said to survive without being identical with each other. This may be possible. But once again, the proposed scheme of thought would only apply to a world which is very different from ours (though, paradoxically, Parfit's intention has been to apply this scheme to our world). A world in which persons frequently undergo fission would not only be different from our world, but also some of our important concepts would have to undergo significant modifications in their application to such a world. For consider the concepts like pride, remorse and horror, for example. As Shoemaker has suggested,<sup>19</sup> I may not feel proud - not, at least, in the usual sense of the term - for a glorious deed I remember (q-remember) having done, since I am only one of the many offshoots of the person who did it. Similarly, for the feeling of remorse for a cruel deed I remember doing. So also, I may not feel horrified - not at least in the

usual sense - at the apprehension of horrible things to happen to my future selves (if, that is, I am about to undergo a fission). Parfit of course, wanted to make some change in the concept of personal identity; but he never intended - nor would he like the idea - that most or many of our important concepts should also undergo such serious changes. Yet this, as we see, is the inevitable outcome of his proposed scheme of thought.\*

Next, the questions of moral responsibility obviously presuppose personal identity (and Parfit has later shown<sup>20</sup> that with the adoption of a particular theory of personal identity, there would be a consequent change in our standards of morality) and it is hard to imagine how they can be freed of this presupposition. For persons are the bearers of moral responsibilities. Moral judgments are pronounced upon the conducts of persons. It is on the merits or demerits of the persons that reward or punishments are given. But it would be (morally) unfair to punish (or even reward) a person for an action unless he is the same person who did that action. This is important, particularly, because there is always a time gap between the commission of a crime and the passing of moral judgment on it, so that the possibility is not ruled out that the person being tried and the person who committed the crime may not be the same person. Hence the importance, in these doubtful cases at least, of first determining the criteria of personal identity before moral judgments could be pronounced, the fact of personal identity being otherwise taken for granted in normal circumstances. Thus Locke: "In this personal identity is founded all the right and justice of reward and punishment."<sup>21</sup> This should pretty well convince one that as long as moral questions are considered important, questions of personal identity must also be important, and

---

\*It will be shown in the next chapter that this sort of conceptual revisionism is the result of a fundamentally misguided idea that the so called 'problem cases' create a conceptual difficulty to the issue of personal identity.

that the former would inescapably presuppose the latter. Without a continuing person, ascriptions of moral responsibilities would be pointless. In his first article "Personal Identity" (PR,71), where he raises the problem, Parfit only claims that questions of moral responsibility could be freed from the presupposition of the questions of personal identity, but he never showed how this separation could be effected nor has he shown how moral questions can still be raised after personal identity has lost its importance. However, he has suggested elsewhere<sup>23</sup> as to how this can be possible. He says, "we may, when thinking morally, focus less upon the person, the subject of experience and ... more upon the experiences themselves".<sup>22</sup> Now, in the first place, this fact of focussing less on the person does not eliminate the person (and so his identity) altogether. If, however, Parfit would go to the extent of 'focussing' solely upon the experience to the total elimination of the subject of experience, it would be difficult to understand why - if not how - moral judgements should be pronounced at all. Presumably, a moral judgment may be pronounced in some such way: action A is blameworthy, and deserves punishment, but this would be entirely pointless if there is no way of giving effect to the moral judgement, and it can be given effect to only if the wrongdoer is identifiable. Thus it becomes amply evident that questions of personal identity are presupposed by the questions of responsibility and the importance of the former remains undiminished as long as the latter are considered important. Perhaps it is also clear that questions of moral responsibility are important - that moral responsibilities are ascribed at all because there is a concept of personal identity. In fact, one can say that without the latter there would be no need of the former. But, strangely enough, Parfit's suggestion points to the contrary.

PART - III

## Chapter 4

### THE TWO CRITERIA

#### Prologue

It has been customary in the discussions of personal identity to talk of two fold criteria of personal identity: the bodily criterion and the memory criterion. Admittedly this is an overschematic account of the criteria we use in making personal identity judgments; but in actual fact these are the two main criteria that are used, and any other criteria that may, in practice, be used can be seen only as special ramifications of, and subservient to, the one or the other. For example, besides these two criteria, we may justifiably work upon certain other factors like blood tests, finger prints and the like on the physical side, and character, personality, skill etc on the psychological side. As regards the former it can be seen without much difficulty that if criteria they are, and are used as such, that is because more often than not they amount to, or point in the direction of, bodily continuity. To be sure the same blood group or similar finger prints do not imply that the present subject is the same person as the earlier person in question - not even that they have the same body; but surely, if no other alternative factor points to the contrary, the said evidence(s) will justify us in saying that here we have the same body and so the same person. But for the usual presupposition (or rather presumption) of bodily continuity, these rather subsidiary criteria could not be used as independent evidence of personal identity, for if they were, it would point to the absurd consequence that a person with the same blood group or with an exactly similar shaped thumb as mine is the same person as myself. As regards the latter criteria on the mental or psychological side, an analogous straightforward reduction of these into the memory continuity criterion is perhaps not an easy enterprise. That is, it cannot be said that similarity of character, skill etc show, or even empirically indicate, continuity of memory. But it will be generally admitted that similarity in these respects would be entirely worthless as evidence of personal identity if it did not carry with it the general

presumption of ~~member~~ memory continuity; for if, in any case, we are sure that a later person remembers what an earlier person did then these other considerations are needless, and if we are unsure of the "memories" being real remembering, these considerations will only serve as mere corroborative facts. (See Penelhum, *Survival and Disembodied Existence*, Routledge and Kegan Paul, London, 1970 pp 60-61) Indeed, more often than not, similarity of character, personality etc. go hand in hand with the continuity of memories; it is, if at all, a very rare coincidence that a person displays striking similarities in these respects while having no memory at all of his past. In a sense, it seems inconceivable (in a weaker sense of course) that this should happen. For it is <sup>a</sup> general fact of nature that our character and personality are largely a product of what we believe, and a large part of what we believe is constituted by what we remember about our past. e.g. a person is likely to be of a contented disposition if he remembers and believes that certain good things have happened to him. There are even some particular aspects of human character that are dependent on what they remember and believe; e.g. a person is revengeful if he remembers and believes that certain acts of cruelty and sheer injustice have been done to him or to people he loves or cares for. To be sure, this dependence will not be logical but only empirical or contingent, but it is the latter, and not the former, that we have seen to be required for the notion of criterion. Thus any closeness of character and personality etc. that would enter seriously into identity considerations will be of a sort that would strongly indicate - if not "entail" (pace Penelhum, *Ibid* p 61) - a fair amount of memory continuity. It follows, therefore, from what has been said so far that the bodily criterion and the memory criterion are the two main or basic criteria of personal identity; and even if anyone doubts or denies what I have said about character and personality etc, he will have little dispute with this conclusion about bodily continuity and memory. He will, however, dispute the latter if he

can show (which I suspect he can't) that similarity of character, personality etc is itself as good a criterion as any of these two. Moreover, that these two are the main criteria, and not either of the subsidiary groups, is evident from the fact that we make fairly confident and reasonable judgments of personal identity on the basis of memory continuity with or without the similarity of character and personality etc, and that we make such judgments on the basis of bodily continuity without needing to look for blood tests and finger prints etc. It is important to note that the facts do not point equally in the reverse order.

With these forewords let us now attend to a reassessment of the nature and status of the two criteria.

#### Section (1): bodily continuity re-examined

The importance of bodily continuity as a justified evidence of personal identity is due, certainly to the fact that persons are known as embodied beings. We know of persons and learn the meaning of "person" by being shown these flesh-and-blood beings and by coming into contact with them. Any other idea of persons that, e.g., they are essentially non-physical in nature and could be completely non-embodied and so on, will at best be an extension of the ordinary concept of a person (in fact we shall shortly be able to see that the idea of a purely non-embodied person, in the Cartesian sense, is not even intelligible). But the fact is undeniable that unless persons had bodies no one could learn, and none could be taught, the meaning of "person". It is this fact about the crucial role of the body involved in the concept of a person that supplies a strong rationale for the sameness of body being criterial for personal identity. For as Locke and Shoemaker have pointed out, what you use as criteria of identity depends largely on the nature of what it is that is being judged.<sup>1</sup> However, even if persons have to have bodies, it may not

seem too obvious that they have to have the same body in order to be the same person. The following comparison might tend to strengthen this scepticism: It is essential for the concept of a parliament that there must be a definite minimum number of members; but, it might be argued, for the continuance of the parliament there need not be the continuity of its members; many and even all of the members may be replaced by new members, and yet the British Parliament is not for that matter replaced by some other parliament or some other thing. If this argument is sound it will purport to show that what is essential for something to be an x need not be necessarily the same in order for the x to be the same x. But I don't think that the argument is sound. For it conceals a crucial confusion between (a) there being members is essential to there being a parliament and (b) there being some particular members is essential to there being a parliament. The fact is that it is (a) that is being claimed when it is claimed that there must be a definite number of members in order for there to be a parliament, and the continuity of this (essential) feature is unhinderedly built into the criteria of parliament - identity. The discontinuity of particular members does not, as <sup>it</sup> is not supposed to, affect the last for, as we have made it clear, (b) above is not an essential feature of there being a parliament; all that essentially is required is that there must, at any moment of the parliament's history, be the required minimum number of members. This continuity-of-members condition is maintained despite the discontinuity of individual members and through the latter's replacement by other, new members; indeed, as we have already seen (Ch 1 sec 1), it is in this last sense that the continuity through time, of the spatial things is to be understood. Now, what is thus true of parliament is analogously true of persons, in so far at least as both are spatial things. So if having a body is essential to

being a person (in the way in which we have claimed it to be), then bodily continuity has got to be a criterion of personal identity. Indeed if having a body can be shown to be a necessary condition of being a person it will accordingly have a strong tendency to make bodily continuity a necessary condition of personal identity. Although I am inclined to think that this is so and have made it considerably plausible in the last chapter, my concern will not be to show that it is so. For what I am interested in is to show only that bodily continuity is a criterion of personal identity which, in my view, need not be a necessary condition. What I claim to have shown so far is that since having a body seems crucial to the understanding<sup>of</sup> the concept of a person, bodily continuity has very naturally been a criterion - and an important criterion, as I show later on. And as in the case of parliaments so in the case of persons, bodily continuity has to be analyzed in the same sense as the continuity of members: nothing in a body needs to remain (literally) unchanged in order for the body to be the same body over a period of time so as to account for the sameness of a person; all that is required is that at every moment of the person's existence there should be a body that is substantially<sup>2</sup> the same as its immediate neighbours. Once this condition is satisfied with respect to the bodies of persons that will be a justified evidence of personal identity. With definite knowledge of bodily discontinuity and in the absence of any alternative criteria, we will have no justification for saying that any pair of non-contemporaneous persons are the same.

We have said before that the notion of a criterion has a unique intermediary logical status in the sense that although a criterio-logical relation cannot be equated with the logical relation of entailment, yet it is not for that reason an empirical or contingent relation. We had claimed that this privileged status is that of a necessary of conceptual

relation earned by "criteria" from the way certain concepts are formed and the way their meanings are learned. It is for this reason that we put pressure on the concept of a person to bring out the importance of bodily continuity as a criterion of personal identity. But even as a criterion, bodily continuity has assumed much greater importance than its rival, the memory criterion. Many writers on personal identity - notably Williams - have underlined this fact by claiming that bodily continuity is a necessary condition of personal identity. By contrast, very few in current philosophical circle have wanted to make similar claims about the memory criterion, and its status even as a sufficient condition has come under severe fire.<sup>3</sup> Let us explain the naturalness and plausibility of this state of affairs by putting some further pressure on the concepts of a person and personal identity. We have indicated above the crucial role of the body in our coming to know persons and learning the meaning of "person". We also know of persons as having certain personality and character and, above all, as being capable of remembering a considerable amount of their pasts. But it cannot be said that the latter set of features are equally crucial - if not more so - to the concept of a person as having a body. For our knowing something before us as a person is rarely, if ever, affected by the suggestion that he does not, or even cannot, remember anything of his past; at any rate, knowing that x remembers or can remember certain things of his past does not seem to form any very essential prerequisite of one's knowing that x is a person, and of applying the concept "person" to x. Similarly, though with less ease, it may be suggested to us that someone, y, before us has no feeling at all - and not any determinate attitude towards things and people: at best his character and personality can be described as one of literal indifference;\*

---

\*If it be objected that literal indifference is still a sort of character and signifies a sort of personality, all I can say is that this is like saying that a blown out flame is still a flame!

but this will have little tendency to give a jolt to our knowing him to be a person, nor for that matter are we likely to call him a degenerated person. But, by contrast, if an x before us does not have a body of the appropriate sort we will be at a loss to call him a person despite his (or its) displaying all or most other characteristics of a person. (A person with no body of any sort would be even harder to understand; at any rate, as we have seen, knowing that something is a bodied being forms an essential precondition of our knowing that he is a person.) It is this difference in the importance of the two sets of characteristics of persons that is also reflected in our normal talk of persons. We will most naturally say of somebody that he is a completely changed personality and/or he has no memories of his past - the "he" undoubtedly referring to one and the same person. On the contrary, in the absence of bodily continuity, no amount of accuracy in "remembering" the past of some person and/or no amount of similarity of character and personality etc will have any justifiable tendency in us to say that this person is the same person as the earlier person. We shall be certainly amazed by the striking accuracy and similarity of these psychological features and at best we would seek the help of philosophical arguments to show that it is nonetheless a case of personal identity.

I am not suggesting that persons are mere material objects - not even that we know them as such. Indeed, we have already distinguished persons from mere material objects and even from the lower animals by virtue of the fact that they are capable of certain "higher order" activities of consciousness. But this is not incompatible with saying that we know them basically as material objects. Of course, this material object, we have been insisting, has to be a body of the appropriate kind - it is what we see as a person, and seeing it as a person, obviously includes seeing it as having consciousness of a very special kind which enables us to know him as a person as distinct from the lower animals. But knowing persons in

the last respect, which of course is very essential to our learning the meaning of "person", requires that we have learnt to, or have mastered the technique of, interpreting some of their bodily behaviour as intentional behaviour, and there is no other way in which we would be able to do so except by observing their bodies. Even we cannot understand fictitious persons, like Hamlet, as persons (even as fictitious persons) except as their having a minimum of bodily features. This shows the crucial importance - indeed the feeling of essentialness - of having a body to the concept of a person. And if having bodies is thus important to being persons then it is but natural to insist, as we do, that the sameness of a person must lie in the sameness of the body. It is equally natural, in view of the foregoing, that for this purpose the sameness of the psychological features does not matter as much - although it does matter. I would suggest further that the last can be explained yet again by appealing to our logical or linguistic intuitions about the notions of sameness and continuity. For it is admittedly less natural to talk of the "same memory and character etc" than it is to talk of the "same body". (This goes also for continuity as applied to both contexts.) All we can mean by "same memory and character" is similar or exactly similar memory and character, for there can be nothing continuing in this case, even in the usable sense of "continuity" as we have specified in the case of bodily continuity, to account for the ascription of sameness. It is for this inevitable reason that "same character" and "exactly similar character" (the same is true of "memory") are often used interchangeably, whereas "same body" and "exactly similar body" do mark a real difference.<sup>4</sup> Elsewhere ... earlier, ... it has been suggested that even if we try to articulate a sense of "same memory" in order to spell out an analogous difference this can be achieved only by making this "same memory" explainable in terms of "same body" - which would point to the primacy of

the bodily criterion anyway. However that will be a matter of philosophical interpretation which has little to do with our linguistic intuitions.

We have thus seen that our logical or linguistic intuitions about the concept of a person and so of personal identity reveals the significance of the bodily criterion and its natural primacy over the memory criterion of personal identity. It may, of course, be said that what we know when we know the meaning of "person" and "personal identity" is not quite the same thing as what constitutes the meaning of 'person' and 'personal identity'. In a sense we have seen that this latter question is fruitless since not any very useful, non-trivial definition of personal identity can be possible. But at any rate, as our discussion in this section would indicate, the question of how we know the meaning of a concept cannot be separated from the question of what that concept means. As Shoemaker said, "what we mean when we assert something to be the case cannot be different from what we know when we know that thing to be the case"<sup>5</sup>. (my emphasis). And what we know when we know that something is an x is not just any fact about that thing but only those facts which justify our applying the concept of x meaningfully to that thing and without which we will not be able to apply the concept "x" to that sort of thing. For example we may know about an earlier person and a later person that both are white men, that both speak English or that both have short hair and so on, but these facts will have nothing to do with our saying that the two persons are the same unless they can be known to satisfy the requirement of bodily continuity and/or memory continuity. It is this latter fact, and not just any (other) facts we may happen to know about them that makes meaningful\* application of "the same person" possible. (This is yet

---

\*Without learning to apply the concept on the basis of these facts, it will not be possible to use the concept meaningfully. A child, for example, having heard the phrase "same person" for the first time from elders, might keep on repeating it or even say it while pointing to a person; but that will not be a meaningful application of the phrase. But having so learnt to apply the concept, if we apply it without having these bases or anything like these, our application of the concept, though not meaningless, would be unjustified.

another way of explaining the close connection between criteria and meaning.) Thus what we know (in the described sense) when we know or learn the meaning of a concept is what we must know in order that meaningful application of the concept can be possible in the first place and without knowing the former any application of the latter, though not necessarily meaningless, is unjustified. From this, together with what I have said about learning the meaning of 'person' and 'personal identity', it is evident that bodily continuity forms an essential criterion of personal identity and certainly a more fundamental criterion than memory continuity. Indeed, but for our inclination in some unusual circumstances to make personal identity judgments solely on the basis of the psychological criteria (and despite bodily discontinuity), bodily continuity could fare well as a necessary condition of personal identity. In the last chapter (section 11.) we examined such unusual cases (namely the cases of alleged change-of-bodies) under two classifications of what we called the "naive theses" and the "scientific thesis". The former, which draws upon the memory continuity criterion without explaining why at all that there is ~~this~~ continuity, has been seen to be implausible and open to serious logical difficulties. The "scientific theses", because it explains the memory- and-character continuity in terms of the spatio-temporal continuity of the physical basis of memory and character etc, was allowed to be plausible and reasonable; but it is precisely because of this explanation we argued that it betrays its inventor. For what it shows is not that the memory criterion is the sole, or even the primary criterion of personal identity, but rather that bodily criterion is the necessary condition of personal identity since the memory continuity, which in this case is reasonable guide to personal identity, is due to bodily continuity. This analysis has obliged us to modify the concept of bodily continuity in terms of the spatio-temporal continuity of that part of the body which is the physical basis of what we described as the "personal faculties". We

have given adequate reason to justify this modification and to abandon the usual rigid sense of bodily continuity. We will not therefore repeat these arguments ~~here~~ except mentioning that such modification will readily appeal to our logical intuitions about the role of body as it relates to our understanding of the concept of a person. For it will surely be admitted in our common discourse about persons and personal identity that even if a person loses a hand or a leg or a kidney, he will still be a person and the same person as the person before this change; but our judgments of personal identity will be considerably affected if he loses his brain (and with it his "personal faculties"). Plastic surgery, kidney transplants, or even heart transplants will not make us say that someone has become a different person, but a brain transplant with the consequent transfer of the memory and character etc will give us very strong reason to say that we have here a different person. This speaks reasonably enough for our analysis of the "Scientific Theses" and our modified version of the bodily continuity criterion. In what follows I shall try to expose the limitation of the "naive theses" from an entirely different point of view.

It has been made amply clear by now that our understanding of persons is basically as a category of spatio-temporal objects. Accordingly, without the spatio-temporal continuity of anything it will be impossible to know what personal identity is and hard to imagine what it would be like. This is not to say that we have to directly observe spatio-temporal continuity of a human body whenever we make, or are prepared to make personal identity judgment; but whenever we do make such judgments in the absence of direct observations of spatio-temporal continuity of the relevant kind, we presume that whatever other criteria we use in making these statements are evidence(s) of, are due to, such spatio-temporal continuity. We have seen before (see esp Ch 3 sec 11) that if remembering is taken as an evidence of personal identity it is because we take the

memory claim to be real remembering which, in turn, suggests the spatio-temporal continuity of the person i.e. of the living human body. Now, to revert once again to the question of how we learn the meaning of a certain concept, we are taught the meaning of "remember" by being told (or shown cases) that the person himself did what he now says he remembers i.e. the person making the memory claim now is the same person as the earlier person who did what is being "remembered", and without begging any question, the last can be explained to us only by telling (or showing) us that the rememberer and the doer of the action are bodily continuous. If this is so it becomes amply clear that identity in terms of bodily continuity is our primary concept of personal identity, and if so, any other idea of personal identity can make sense only by depending on this primary sense, but not conversely. Now, if the "naive thesis" of alleged change of body does make sense and does seem to be plausible case of personal identity without there being bodily continuity, that is because (and that's a conceptual "because") there is this primary sense of personal identity. This conceptual dependence can be explained by the fact that if there were not persons who are bodily continuous with other non-contemporaneous persons, we would not be able to apply the concept of same person in the first place, and the "naive thesis" which makes use of the notion of 'same person' only in secondary sense derived from the primary sense, would not be an intelligible thesis at all. This is due to the fact that there will be the secondary use of "same person" only if there is the primary use of the concept and not conversely. This interesting distinction between the primary use and the secondary use of our concepts I owe to Professor Norman Malcolm through one of his unpublished(?) papers.<sup>6</sup> As he put it in that paper, just as we can say "the doll has pain" (and that is only a secondary

use of the concept of pain<sup>7</sup>) so also we can say in some unusual cases\* that here also we have the "same person" (which, Malcolm argued, would be equally a secondary use of "same person"). And he explained this by saying that just as we can say the former only because we have the primary use of "pain" in the human context (but not vice versa), so also we can say the latter only because there is the primary use of "same person" in the normal bodily context (but not vice versa). Malcolm used this argument against Descartes to show that just because there is this secondary use in which we could still talk of the same person and so of person even in the known absence of a body, Descartes was misled to reason on this ground that person is essentially non-physical (a "thinking thing"). But if Descartes were right on this, it would follow that all the persons in the entire universe could have been non-embodied (as distinct from dis-embodied) thinking things, yet, as Malcolm's argument shows, this would make no sense since the latter, being only a secondary use of "person", is dependent (conceptually dependent, I claim) on the primary use of "person" as signifying embodied beings. This argument, which I think is basically sound, I also take to imply that just because we can, in a secondary sense, talk of "same person" in the cases of the described sort (see the last footnote), it does not follow that it is a genuine case of personal identity and that personal identity could have been only of this non-physical type. For the latter being only a secondary use of "same

---

\*The unusual cases he imagined come safely under my classification of "naive theses". For he imagines few cases like the following: a person Robinson, is dead and buried; later a voice is heard by friends and relatives that sounds exactly like that of late Robinson, and the voice reports and "remembers" what only Robinson could have known and thus reported. In the face of such baffling accuracy in memory claims and in view of our general reliance on memory as a reasonable guide to personal identity, we may say that we talked to Robinson - thus using the concept of same person. But, Malcolm argued, this would only be a secondary use of the concept.

person" is conceptually dependent on, and owes its intelligibility to, its primary use which, if what we have said is right, goes hand in hand with bodily continuity and without this the secondary use of the concept in itself would not make any intelligible sense. This latter will have to be the fate of the "naive thesis" if it claims to supply the standard case of personal identity with its expressed claim that memory is the sole criterion of personal identity and the implied contention that persons are essentially non-physical thinking things. I am not sure that Malcolm would agree to my last conclusion; but I do not see how then he would consistently maintain his position against Descartes, namely that the Cartesian idea of pure non-embodied person is unintelligible. I want to maintain further that any theory that purports to stress the non-physical nature of persons and to establish that the criteria of personal identity are solely non-physical is guilty of taking the secondary use of these concepts seriously and of trying to give it the status of primacy. In so far as the "naive thesis" purports to do this, it must be mistaken. These theories delude themselves into thinking that we are talking about the same thing here as we do in the normal cases presumably on the superficial ground that in the former case our familiar words like "person" and "same person" are used in their familiar configuration. It is because of the latter that what they say would not be straightforwardly nonsense, but in being committed to an entirely non-physical concept of person and personal identity and as such abandoning the primary sense of these concepts, they would not be describing what is the case. Wittgenstein expressed even<sup>a</sup> stronger view in his following remark on the secondary use of our concepts, "... the fairy tale (in which e.g. a pot can be said to see, hear or even talk, PKM) only invents what is not the case; it does not talk nonsense".<sup>B</sup> (First emphasis added).

I shall argue later on that the secondary use of "same person" in the "naive" form of the problem cases, gives a degenerated sense of this concept.

However, what will not be disputed, and what will suffice for my present purpose, is that this type of use of our concept of personal identity cannot be justified. Malcolm's ground for saying that they are only secondary use of "person" and "same person" was that we can not use the bodily continuity criterion in these cases. And if my assessment so far of the memory criterion is correct, then in the absence, ex hypothesi, of bodily continuity the memory-criterion is, in principle, inapplicable since, as we have argued, the latter is a criterion only because bodily continuity is a criterion of personal identity (see Ch 3 sec 11 especially).

It follows therefore that the "naive thesis", which because of its apparent plausibility, inclines us to say that the criteria of personal identity are purely non-physical and thus purports to undermine the importance of the bodily criterion, is not strong enough. Through its inherent weakness, we have been able to see the strength of the bodily criterion as an essential criterion of personal identity and certainly as the primary criterion thereof. In the next section we will examine some of the main points of the claim that memory continuity is the sole or even primary criterion of personal identity; and by showing this claim to be mistaken in all its interpretations, we will establish the fact that memory is only a secondary, not an independent, criterion, which has to depend on the bodily criterion for its use.

Section (ii) Memory: not an independent criterion

In the last section we have gone quite some way to argue that the bodily criterion is the primary criterion of personal identity; we also argued to the effect that although in most cases bodily continuity and memory continuity are used to determine personal identity yet in certain cases bodily identity can be used, and is used, as the sole criterion. By this I mean that even if in certain cases there is considerable discontinuity (dissimilarity) of memory, character and personality, etc., yet personal identity judgment can be made, and is made, if bodily identity is assured. But I suspect there is an ambiguity involved in the claim that a certain criterion (and this applies to the memory criterion as well as to the bodily criterion) is the sole criterion of personal identity, which I must make clear before going to consider this claim in respect of either or both of the criteria of personal identity. The ambiguity (as applied to the bodily criterion) is between this sense just outlined and the sense in which the claim amounts to saying that bodily identity is the only criterion that is used and can be used in determining personal identity. Similarly, the claim that memory is the sole\* criterion of personal identity can mean either that in certain cases, even if bodily identity is known to be absent, identity judgments can be made on the basis of memory or that memory is the only criterion that is used and can be used to determine the identity of persons at different times. It seems to me that no clear distinction has been made between

---

\*Indeed, when the nature of this claim is exposed in the two senses, as I have exposed it, to say that a criterion is the sole criterion in sense 1 will be a very odd thing to say. As Professor Swinburne suggests to me, the proper thing to say here would be "a sole criterion". But although this is true and is in the spirit of my distinction, yet this description again would sound very unnatural and monotonous - if not odd. In view of this, as well as in view of my intention to expose the mistakenness of the claim (in respect of memory), I shall continue to state the claim in its familiar wordings, even in its sense 1, since that is how it has been made - although I am in no fear of being committed to the oddity mentioned above in view of the way I have stated the distinction between the two senses of the claim.

these two senses of the claim - let us call them sense 1 and sense 2 respectively - in most writings on the subject, although each of the claims that bodily identity and memory is the sole criterion has been quite often debated. And rather strangely, (so at least it seems to me) the answers given to each of the claims have been given<sup>at</sup> two different levels (without ever realising that this has been so). For usually when it has been claimed that memory is the sole criterion of personal identity this is claimed in sense 1 whereas when it is denied that bodily continuity can be the sole criterion of personal identity it is denied in sense 2.<sup>1</sup>

This becomes clear if we consider the basis on which such claims are usually made (or denied) in each case. For the source of the idea that memory is the sole criterion of personal identity has been the plausibility of the so-called puzzle cases where either someone allegedly turns up with the memory claims and character similarity etc of another person known to have lived in remote past or where two contemporary persons allegedly come out with the memory claims and character etc of each other. It is argued that since, in such cases, it not only makes sense to say, but we are so much inclined to say, that it is a case of reincarnation or bodily exchange (as the case may be), memory must be the sole criterion of personal identity. I am pretty sure that people who argue along this line will not say that memory is the only criterion that we use whenever we say that a later person is the same person as an earlier person. And even if they do want to argue to this effect, my argument will be to point out that this does not follow from its purported premiss, and that their claim would certainly need more argument than simply pointing out what we would say in the puzzle cases. But given things as they are, it will be enough to point out that although in certain cases we may be inclined to say that memory is the sole criterion of personal identity, <sup>we will not say that it is the</sup> only criterion that we can use in making personal identity judgment; for given a legitimate sense of "criterion" there are hosts of other cases where we

do use the bodily criterion besides, and despite the absence of, the memory criterion in making personal identity judgments. I take it, therefore, that the only sense in which it can be claimed, and I suspect it is the sense in which it has usually been claimed, that memory is the sole criterion of personal identity in the sense 1 delineated above. Now as regards the parallel claims that bodily identity is the sole criterion of personal identity, very rarely, if ever, it has been claimed explicitly - though its primacy and even necessity has been often argued for -, but I have seen this claim expressedly denied. And when it is denied that bodily continuity can be the sole criterion of personal identity it is denied in (our) sense 2, for it has been denied on the ground that there is another criterion - the memory criterion - which is equally a "non-inductive evidence"<sup>2</sup> that justified us in making identity judgments on its basis, and that this clearly means that bodily identity can't be the only criterion of personal identity. I suggest that saying this is not incompatible with saying, as I do and as many would be willing to do, that bodily criterion is the sole criterion in sense 1, for as I said, clearly there are many occasions when we do say that a later person is the same person as an earlier person because they have the same body - even if they are very different in respect of memory and character. But it seems to me that although many would not dispute the last description of the situation, they would not - at any rate did not - take this to be expressing the claim that bodily identity is the sole criterion of personal identity. They would, on the contrary take this claim only in sense 2. Shoemaker<sup>3</sup> again. He, for example, gives an account (which he eventually rejects) of this claim that is clearly on the lines of sense 2, and not of sense 1. For he implies that if bodily identity were the sole criterion of personal identity, there would have to be no other criterion and that it would "rest on a mistaken view" to think that there are; and he also says, quite rightly, that if this were the status of the bodily

criterion then whenever we took a person's memory claims as evidence of his identity through time, we should be doing so by inductively inferring it as reliable evidence from facts about bodily continuity.\* Now, although this is not true<sup>†</sup> and bodily identity cannot thus be the sole criterion of personal identity yet it is, as I have shown, the sole criterion in another sense (i.e. in our sense 1); for if this latter sense were not there it would not be possible even to claim that memory is the sole criterion since, as I have pointed out, this claim has been made precisely in this latter sense. It is important, therefore, to make the distinction, which we have made, between the two senses as above, otherwise we would be arguing somewhat at cross purposes if we argue that memory is the sole criterion and bodily continuity is not the sole criterion of personal identity. I suggest that it is this lack of explicit distinction between the two senses (which nevertheless is there and is at the root of the above line of reasoning about the two criteria) which, partly at least, has misled many into thinking that the memory criterion is more important than the bodily criterion in the way it has been thought to be. I have so far been able to uncover at least one such mistake that has been elevated to the status of a theory. And this, it seems to me, is the thesis suggested by T.E. Wilkerson in his Minds, Brains and People that "mental identity is sufficient (but not necessary) for personal identity; bodily identity is both defeasibly sufficient and defeasibly necessary for personal identity".<sup>4</sup> My concern is over the qualification "defeasible",

---

\*i.e. we would discover empirically, using bodily identity as our criterion of personal identity, that the sincere and confident memory-statements are generally true, and since this present case seems to be one such case of sincere and confident claim, we would infer that this may be a case of personal identity.

<sup>†</sup>This is not true because memory is used as a criterion, not by virtue of being an inductive evidence of personal identity but by virtue of a logical connection between the concepts of memory and personal identity.

for even as a sufficient condition of personal identity this qualification purports to endow the memory criterion with an importance which tends to give it a superiority over the bodily identity criterion even as a defeasibly necessary condition of personal identity, since the latter can be allegedly defeated in certain circumstances whereas the former can never be (so the thesis tends to claim). The reason why I think it to be typical of the error outlined above is the following: the thesis holds that the bodily criterion is defeasible because in certain circumstances it can be defeated by the memory criterion, and this is but an echo of the claim that bodily identity cannot be the sole criterion in our sense 2. On the other hand the thesis holds that the memory criterion is not defeasible in this sense, and the only reason that seems to justify the last claim is that under certain circumstances (i.e. the alleged cases of bodily exchange) memory seems to be the sole criterion and does not get defeated even by the known discontinuity of bodies. And this only means that (provided that this explanation of these situations is correct and reasonable which, I will presently show, it is not) memory is the sole criterion in my sense 1. Thus, without realising that there are two different senses in ONE of which bodily continuity is not the sole criterion and in ANOTHER of which memory is the sole criterion, the thesis under examination seems to be taking the claims that bodily identity is not and that memory is the sole criterion simply at their face value, and on this basis gives mental identity an undue importance over bodily identity. It is certainly an undue importance because, for one thing, it cannot be held that the memory criterion is not defeasible under any circumstances. As an interesting first step towards seeing the falsity of the latter we have only to recall the fact that in most cases of day-to-day life memory discontinuity is defeated by bodily continuity (and this most certainly makes memory ineligible as a necessary condition of

personal identity), and to see that, as a sufficient condition, memory continuity is defeated by bodily discontinuity, we have to imagine certain slightly abnormal or unusual cases like Williams' memorable example of Charles/Guy Fawkes which is a clear case in point. Paradoxically for him, and much to our delight, Wilkerson states with apparent approval<sup>5</sup> that "our reluctance to say that Charles and Guy Fawkes are the same person is due to our implicitly insisting that they must at least share the same body to count as the same person" (Id p 27). And if this is so memory is as much a defeasible sufficient condition of personal identity as bodily identity is, and so can't be the sole criterion in sense 2. Yet because of the fact that it seems to be the sole criterion in sense 1 (i.e. in the alleged bodily exchange cases like Mr and Master Bultitude case), coupled with an oversimplified view of a criterion being the sole criterion of personal identity, it has been mistakenly thought that memory is the more important criterion, and has been, consciously or unconsciously, kept untouched by defeasibility. Incidentally, I suspect that some more normal situations can be brought in to show our point that memory continuity is as much defeasible as bodily continuity is. Quinton considers such an example of a pair of absolutely identical twins<sup>6</sup> whose characters and memories are totally indistinguishable and "whose thoughts and feelings have been precisely the same since the first dawning of consciousness in them". Much to the embarrassment of his own theory (that personal identity is constituted by mental identity), Quinton concedes that the later phases of twin No 1 would be as much continuous, in respect of memory and character, with the earlier phases of twin No 2 as they are with his own earlier phases, and that yet it would be absurd to say that they are the same person, and this will be absurd because they have two different bodies. This would be a clear case of the memory criterion being defeated by the bodily criterion. Quinton, however, tries to avoid this conclusion

by explaining that here "we might ignore the duality of their mental states, but we should be able in principle to assert it", for, he goes on to say, "however alike the characters and memories of twin No 1 on Tuesday and twin No 2 on Wednesday, they will inevitably be less continuous than these of twin No 2 on these two days" (p 405). Now, apart from sounding too trivial (i.e. they must be less continuous because they belong to different persons), this seems to be an extremely forced explanation and a theoretical cover-up. For, in the first place, knowing fully well that it is a case of two persons and so the memory criterion - despite its striking continuity - will not do the trick, Quinton seems to be content with supplying an explanation in terms of memory-and-character continuity anyhow. And, secondly, the memory criterion, as it is supposed to be used as a guide to personal identity, does not require that all (literally all, as Quinton seems to imply\*) memories must be preserved, but only that a later person must have more or less the same memories as the earlier person - or that the later person must be able to remember a considerable amount of what happened to the earlier person - in order for him to be identical with that other, and therefore the sort of memory-and-character continuity, striking as it is in the case of the absolutely identical twins should be enough to enable us to identify a later phase of twin No 1 with an earlier phase of twin No 2 as belonging to the same person. But the fact remains that it fails, and it fails because it is defeated by the known fact of bodily discontinuity. (In a sense there is no need to say the last; since  $p_1$  and  $*p_1$  both satisfy the memory criterion at the same time, the bodily criterion takes over).

---

\*Perhaps, being himself persuaded to think that memory and character etc is the necessary condition of personal identity, Quinton is clinging to such an explanation. But we have already seen, especially in our discussion of Locke and Hume, such a status of the memory criterion can't be maintained.

Now, if what has been said in the foregoing is right then it follows that there are certain cases in which the memory criterion too is defeated and hence that memory cannot be the sole criterion in sense 2. But can it be the sole criterion in our sense 1? As we have seen, it is the apparent truth of this last claim which has been (partly, at least) responsible for the mistaken importance of the memory criterion. My own view is that both memory and bodily continuity are defeasible but reasonable guides to personal identity. (For this is what makes them criteria in my sense of the term). And being so, neither can be the sole criterion of personal identity in sense 2. But whereas bodily identity can be the sole criterion in sense 1 (i.e. in most, though not all, normal circumstances), I feel very doubtful about the parallel status of the memory criterion. To explain my doubts I must consider some very likely cases in which memory has been thought to be the sole criterion (at least in sense 1) of personal identity.

Admittedly, the most likely cases in which we are said to <sup>be</sup> very much inclined to say that memory is the sole criterion (i.e. despite bodily discontinuity, memory can still enable us to make personal identity judgments) are the alleged cases of bodily exchange.\* These cases, we've seen, can be either the Shoemaker-type case of brain transfer (which we have described on the "Scientific thesis") or the Lockean-Quintonian cases of sudden switch of memory and character (which we called the "naive thesis"). For reasons already adduced with reasonable adequacy we may leave aside any fresh consideration of the former type of cases since they

---

\*Perhaps the Charles/Guy Fawkes-like cases, though they are also thought to be lending support to the same supposition, will be considered as not being as strong as these cases of "bodily exchange". Perhaps also, in view of Williams' well-known attack on the coherence of such cases, many recent sympathisers of the said supposition may feel indifferent towards these cases. However, what I will say against the plausibility of the alleged bodily-exchange cases will also apply against such cases; and for this reason I will not discuss these cases separately.

do not show that memory, even in such cases, is the sole or even the primary criterion of personal identity (See Ch 3 section). It is, therefore, the latter type of cases which, if they are reasonable cases of bodily exchange, would show that memory in such cases is the only criterion since ex hypothesi there is no part of the body which is exchanged along with the exchange of the psychological features. A host of puzzle cases have filled the pages of past as well as current literature to lend support to the plausibility of the thesis under examination. The following will give a schematic and neutral account of what would be the case if such a case occurred: to start with we have two persons A and B, with every idiosyncratic differences in respect of physical and psychological features. All of a sudden, from what originally was the body of A there begin to emerge the memory claims and the display of character and personality\* which originally was associated with B, and which even now we will have reason to attribute to B, and vice versa; and this situation continues for a fairly long time, perhaps ever after. In consequence we now have two "new" persons: the A-body-person with what may be called the "B-ish" character and the B-body-person with the "A-ish" character. There is no explanation whatsoever of how this happened - only that it just happened. Now if such things actually occur in our world, and if we have to say something in these circumstances, there are two things, equally plausible but mutually inconsistent, which can be said: that A and B have exchanged bodies or that they have switched memories and character. But if we say the former, we will be working solely on the memory criterion to the total exclusion of the bodily criterion, and

---

\*The question of whether these latter "psychological" features could be perfectly displayed in a different and perhaps very dissimilar body is of course highly debatable, since certain aspects of human character and personality are highly "body-specific" and could not be displayed without an appropriate body. But without going to that detail we can work upon the assumption that they can be reasonably displayed. A fuller discussion of this question, however, has been given in our section on Williams.

if we say the latter we will be working solely on the bodily criterion totally disregarding the memory criterion. Prima facie there will be no reasonable ground to say the one rather than the other. So, in principle such cases have nothing to show that either criterion is the sole criterion (even in our sense 1). But being thus supplied with no logical equipments to deal with such cases the authors of these cases have usually appealed to our logical intuitions, to what we would want to say if such things occurred. But I think that what we would want to say is a very difficult question - especially in extraordinary situations like these: perhaps we would not be able to say anything and our concept of personal identity would break down, or perhaps, as looks very natural and as I shall presently try to show, we will be largely divided over the issue. Yet what seems to be disturbing is that the authors of these imaginary cases have often taken it for granted that it would be easily pronounced as a case of change-of-body. This is suggested by the fact that with perhaps very few exceptions, like Williams' cautious account in his "The Self and the Future", the authors of these cases have presented them, in not so neutral a manner but somewhat contentiously as cases of bodily exchange. The Cartesian picture of a person is often unguardedly exposed as working behind the formulations of such cases.<sup>7</sup> It will appear very likely that being persuaded to think in these lines they have in a way assumed that the only reasonable thing to say in these circumstances will be that which will suit the "bodily exchange" hypothesis, and that the reasons which they usually wanted to adduce have been more suited to this purpose rather than having any rigorous logical force. I shall try to prove my point, first, by reporting what Quinton, a notable and representative contender of the supposed theory, says about what is reasonable to say in these circumstances; and secondly, by examining what he gives as reasons or "supports" for his thesis. After describing a (similar) case of what he calls a "psychophysical" exchange, he asks what it would be reasonable

to say in this circumstance and says that it seems to him Quite clear that we should not say that B and C had switched character and memories" (my emphasis), and goes on immediately to the conclusion that "if this is correct (without showing any reason why this is correct PKM), it follows that bodily identity is not a logically complete criterion of personal identity". (p 602). There is no apparent reason, and nothing is supplied to us yet, as to why can't we say that and why must we say that they have changed bodies - except for an implicit reliance on an essentially non-physical theory of persons (as indeed he says later on that the "soul is what a person fundamentally is").

In the following paragraph, Quinton goes on to produce some "supports" for his professed theory; and the first support for this comes from the "rather weak" evidence (Quinton's own expression) of imaginative literature, where the author shows "not the smallest trace of hesitation" in calling it a case of change-of-body (Quinton here refers to the Bultitude case of Anstey's Vice Versa). The evidence is very weak indeed, for certainly "no hesitation" on the part of the author is no reason for showing or saying that it is a case of bodily exchange. To make matters worse, it can be said that the author shows no hesitation presumably because, like Quinton himself, he had designed or meant this story to act as a persuasive case of "bodily exchange". Second, a "solider support" for the thesis comes from the consideration of how the relatives and friends of the victims of this psychophysical transfer will feel, and what they would like to say. In matters of these Quinton believes that it is the psychological features of persons that matter most - indeed, as he seems to suggest, only these things matter. "In our general relations with other human beings". he says, "their bodies are for the most part in trinsically unimportant" and are only "convenient recognition devices" (p 402, my emphasis). And on these grounds he

concludes that "the soul, defined as a series of mental states connected by continuity of character and memory, is the essential constituent of personal identity" (p 403). This is still a very weak explanation. For easy counter-examples are readily available by appealing to the same logical intuitions. If I am to undergo a radical, and even disagreeable, change of memories and character my mother, I am pretty sure, will have no hesitation in treating me as her son - though of course she would not treat me as tenderly as she used to, but that is a different matter altogether. In the event of this happening my mother will be awfully sorry that evil has befallen her son and my wife will curse her ill luck; but it seems very unlikely that they will abandon me and (still more unlikely) accept an imposter(?), who appears with my old memories and character etc claiming to be their son and husband. Even if they would be willing or rather forced to do something like that (perhaps in the event of my treating them brutally) something like a legal procedure of "abandonment" and "divorce" will have to be brought in, and what is more to the point, they would be doing so, not because they believe that I have changed my old body and gone to "live" in another body but because they find my changed character and personality etc intolerable. Further, it will not be unnatural to think that even after all this have been legally done, my mother will still angrily protest, and with genuine feeling of maternal love, if my "old" body is tortured or slaughtered. If all this is correct and equally natural and reasonable things to say, then it follows that Quinton has not given adequate reasons to support his thesis. In general the so called inclination to say in such cases that a bodily exchange has occurred has no logical support and even as an inclination, it is not any the more natural as it has been supposed to be. Hence the supposed consequence of this inclination, the theory that memory is the sole criterion, at least in sense 1, seems to be very dubious, to say the very least.

A more substantial criticism of the Lockean-Quintonian type of theories of alleged bodily exchange can be made by showing that it, being only a version of the general memory theory (where all we are to rely on is the mere unsupported memory claims), will be no less open to the logical difficulty of reduplication. Following Williams' lead, we can say not only that the A-body-person will claim to remember what only had happened to B and vice versa, but the following may possibly occur: another person C may simultaneously turn up with the "memories" and character etc of B, so that we would now have two persons claiming identity with one and the same person B; and we have no logical ground for saying that only one of them is, and the other is not, the same person as B. And if this is so, it would be as much "vacuous" - though not meaningless\* - to say that the A-body-person is the same person as B when he alone underwent the described change. The same consideration will make it equally vacuous to say that the B-body-person is the same person as A since, simultaneously, a fourth person D might appear with the "A-ish" memories and character etc. It follows that, since relying on the memory criterion in the said type of cases makes any justifiable assertion (and so ascertainable theory) of personal identity impossible, memory would be worthless as a criterion (= justificatory evidence); and this means that the very purpose of the alleged cases of bodily exchange (some proponents of which even wanted to show with their help, that memory constitutes or makes personal identity) has been defeated by their inherent weakness.

A further weakness of these theories comes out in Quinton's own admission that if the alleged cases of bodily exchange did occur we will

---

\*It should be clear from what has already been said about the possibility of such cases (see especially P 156 above) that making personal identity judgment on the described ground will be logically incoherent if memory is meant as an analysis of personal identity, and even as a criterion of it. A further account of the general philosophical limitations of such theories with special reference to split-brain transplants, will be given in Appendix 2.

have to "extend"<sup>8</sup> the concept of a person and of personal identity and say that a person is where his memories and character are. (Saying this is prima facie ambiguous between 'changing the meaning' of "person" and 'extending the application of the concept'. Quinton did not make it clear what exactly he meant by saying that there will, in the described cases, be an extension of the concept of a person - though apparently he seems to think that there will be a slight change in the meaning of the concept. Saying the last straightaway might raise certain amount of controversy; but what will not be disputed is that if, under these circumstances, we want to say that the person is where his psychological features are, we will be making an extended application of the concept of a person). This brings us back to our distinction, made in the last section, between the "primary" and "secondary" use of our concepts. For if we are thus to extend the concepts of a person and of personal identity, we will be allowed to talk of the same person and people will, perhaps, have no difficulty in understanding what we mean, but that will be a secondary use of "same person, and like Wittgenstein's fairy-tale it would, perhaps, be inventing what is not the case."<sup>9</sup> I.e. we would, in that case, be inventing a new way of talking (meant to suit such strange cases) but this will not guarantee that we would be talking about what is the case. As we have explained in the last section, the secondary use is logically dependent on the primary use of the concept but not conversely; and so it would follow from this, and from what has just been said above, that the memory criterion (when it has the airs of being the sole criterion of personal identity), being subservient only to the secondary use of "same person", would only be a secondary criterion, at best, in relation to the bodily criterion, for we have seen that the latter is an essential element in the primary use of "person" and "personal identity". (This is, thus, yet another way of showing the secondary status of memory as a criterion

of personal identity).

Thus, the outcome of our deliberations in the present section together with those in the last amounts to this: both bodily continuity and memory are the criteria of personal identity and that each being defeasible, under certain circumstances, by the other it cannot be said that either is the sole criterion of personal identity in sense 2, i.e., being the only criterion that can be used to determine personal identity. The mistaken claim that memory could be the sole criterion was made plausible by the seeming plausibility of saying that in certain cases it was the only criterion which we use (i.e. as the sole criterion in sense 1 as we have outlined above). It has been seen that this supposition is doubly mistaken, because firstly even <sup>if</sup> it was reasonable to say that memory was the only criterion that gives correct result in those cases it would be the sole criterion only in sense 1 (ignoring the oddity of saying this), and secondly because it has been shown that it is not reasonable to make even this weaker claim. By contrast, it can be said that bodily continuity is the sole criterion in this latter sense, which is evident from the many normal cases where even in the (known) absence of memory continuity it enables us to make personal identity judgments. Fairly reasonably, this explains the fact that the bodily criterion is the more important and the primary criterion of personal identity.\* It has also been shown, by putting pressure on the concepts of person and personal identity that the bodily criterion is the more fundamental criterion which is involved in our learning and understanding the meaning of "person" and "same person". Any other - i.e. non-physical - account of personal identity, it has been argued, would only be intelligible as a secondary account and would hence

---

\*It is significant to note that we have to imagine unusual, puzzle cases to make plausible the claim that memory is the sole criterion even in sense 1; and by contrast it is not only plausible but also true that in many normal cases bodily identity is the sole criterion in this sense.

conceptually depend on the bodily or physical account of person which is the primary account. And since the secondary account gives an appearance of primacy to the memory criterion the latter must be a dependent criterion in its relation to the bodily criterion. It has further been argued, by analysing the way we come to learn the meaning of "remember" and its cognate, "memory", that the correctness of memory-claims could only be tested and established by reference to the bodily history of the "rememberer". This means that if, in most normal cases, memory (i.e. memory-claim) is used as a criterion of personal identity, that is because we normally take it as a reliable evidence, i.e. as a likely case of correct memory. It seems to follow therefore that the use of memory as a criterion of personal identity implies a covert presumption of bodily continuity. (This may be further reinforced by the obvious fact that if a "rememberer" were known to be not sharing the same body with the past witness of what is "remembered", we would not say that he really remembers - or at any rate, we will regard it as a very doubtful case of remembering.) Of course we do not need to establish that a memory claim is a case of real memory whenever we take one as such and use it as our criterion of personal identity. We simply take it as a reliable evidence and that is because, as Shoemaker has argued (See ch 3 see iii above), it is a necessary truth - or at any rate, a general fact of nature - that people's memory claims are generally true (which, for reasons given earlier, is the same thing as saying that they generally correspond to facts about the past histories of the bodies of those who make these claims). It has become quite clear therefore that memory is not an independent criterion and has to depend on the bodily identity criterion in order for it to be used as a criterion of personal identity. For, as we have seen in the last section and as Shoemaker so rightly said, if bodily continuity were not a criterion of personal identity, nothing else would be an evidence thereof.<sup>10</sup>

Section (iii): Disembodied Persons

In the two preceding sections we tried to show the primacy of the bodily criterion of personal identity and dismissed the claim, often made, that memory is the sole criterion of personal identity. The latter was shown to be decisively false if taken in our sense 2 (see sec ii) in which a criterion is the only one that is used, and can be used, in making and coming to know personal identity judgments; and it was shown to be very dubious and unjustified if it meant (sense 1) that, in some circumstances, memory could be the only criterion. The last was shown to be the case by examining some specimens of these circumstances (i.e. the puzzle cases) in which memory seemed to be the only criterion, and by showing that they were not as plausible as they were supposed to be. But there is yet another area of speculation which gives the claim a reinforced plausibility and this is the widespread belief in disembodied existence of persons. For if persons could be said to exist in disembodied states, and if this would be a plausible thing to say, then the only criterion for the identity of such persons would have to be memory, since obviously in these cases there is either no body or no continuity of body. Therefore in the interest of our account of personal identity and the criteria thereof this belief in disembodied existence needs consideration in some detail.

Now, basically there are two distinct moves involved in the belief in disembodied existence. The first, which is a rather radical move, is that all persons could have been bodyless and that there could quite conceivably have been no bodied persons at all. I will refer to this belief as "non-embodied existence". The second move, which makes a slightly weaker and less radical claim, is that even if persons are in fact bodied beings, each of them could continue to live or survive in a disembodied state after the destruction and decomposition of his body which normally is called death: this state may hereafter be the perpetual

state of the person or it may be terminated from time to time by the person entering into different bodies at different stages of his personal history. I shall use the expression "dis-embodied survival" (with special emphasis on 'dis') or simply "survival" to describe this second belief. As we shall presently see, it is this second belief, more than the first, which is cumbersome to our account of personal identity - and much more so to any exclusive bodily theory - and is more difficult to tackle. (This is because the first belief, in non-embodied existence is more easily dismissed than the second belief, as we shall presently see; see also p 199 above). For this reason, it will be given a more detailed consideration than the first belief. But before going to (and in order to be able to) deal briefly with the first belief, it is significant and useful to note that this belief and the second are the natural offshoots, in the reverse order of statement, of the two distinct strands of the Platonic-Cartesian theory of persons, namely (1) that a person often (but not necessarily always) combines two radically distinct, contingently related, substances - the body or the corporeal substance and the mind or the incorporeal substance, and (2) that it is the second of these substances which is the person. Now, since the first of the above beliefs, the belief in non-embodied existence, is based on (2) the weakness of this belief would seem to be more obvious than that of the belief in dis-embodied survival. This can be seen in the following way: given the least controversial view of "substance", a substance is that which can exist all by itself without having to depend on any other substance or any other thing. And while persons and bodies quite obviously satisfy this requirement, it is highly arguable if minds even can. It seems to follow from this that the equation or identity of persons and minds is unwarranted. Further, this equation is easily seen to be counter-intuitive since it is fairly more natural to say that persons have minds rather than saying that they

are minds (and most certainly, the latter does not follow from the former). But if this is so, the belief in non-embodied existence, <sup>being based upon</sup> this rather dubious equation, becomes all the more dubious and weakened.\*

However, although the above argument leans heavily on my belief that minds are not substances, I will not argue the point here as I do not think it essential to my purpose. (I shall simply content myself with the fact that my belief draws considerable support from our logical intuitions, as indicated above, about what can and what cannot exist by itself).

Instead I shall proceed to supply some further ground to show that the belief in non-embodied existence is unintelligible. In the last two sections - especially in the first of the two - I made the distinction, following the hints from Malcolm and Wittgenstein, between the primary and secondary use of language, and showed that the latter conceptually depends on the former and not vice versa. Further, by analysing the way the concepts of person and personal identity are learned, I showed that it is the bodied idea of persons that constitutes the primary sense of person and that the idea of disembodied person is only a secondary idea which has meaning only on the assumption of there being bodied persons. If this is so, the idea of pure non-embodied existence - which excludes the conceivable existence of bodied persons - is unintelligible. For if there were no bodied persons at all, we could not be taught, and so could not learn, the concept of a person and consequently there would be no concept of a person in the first place, and no concept of disembodied person either. (See sec i above, especially pp 93-94). The idea of non-

---

\*Incidentally, my apparently more 'tolerant' attitude towards the second belief, the belief in disembodied survival should not be misunderstood. For although I do claim that this belief is less obviously weak and less vulnerable (than the first belief), I do not in the least suggest that persons are bodies. All I have in mind, though I do not commit to anything here, is that a theory that seems to equate a substance (person) with a non-substance (mind) is more obviously vulnerable than a theory that may seem to equate a substance (person) with a conjunction of a substance (body) and some properties (a series of mental states, perhaps).

embodied persons which, with fair amount of justice, can be attributed to Descartes, was shown by Malcolm (see sec 1 above) to have been the result of an illicit reasoning, since it was supposedly based on the premiss that in certain unusual circumstances we do talk of person and same person despite our strongly justified belief that no body whatsoever is involved. It may be conceded, as Malcolm in fact did concede, that our talk of person and same person in these circumstances is intelligible; for presumably we know what we are talking about, that it is a case very much like the (actual) case of a person talking to us and like the (actual) case of a person remembering certain events in his past. But equally, we certainly know what sort of "person" and "same person" we are talking about: a somewhat "non-standard" person (and a consequent "non-standard" case of personal identity) which we make sense of only because we know what standard cases of person and personal identity are. But the standard altogether dropped out of discourse, the idea of pure non-embodied existence becomes an utterly unintelligible notion; for here we would not know what exactly we are talking of. To compare a familiar but salutary analogy of Ryle's, there can be false coins only if there were coins made of the proper materials, issued by the proper authority.<sup>1</sup> Could the idea of a counterfeit ever make sense without this last presupposition? We shall talk no more of "non-embodied existence" and must now proceed to consider the possibility of "disembodied survival".

It would hardly be disputed, especially in view of what has been said above, that the belief in survival, the belief that a person, after his bodily death continues in a disembodied state or in different states of disembodied and bodied existence, owes its intelligibility - whatever it may have - to the concept of embodied existence, which we have seen to be the primary concept. And this acts as a point in its favour as distinct from the idea of non-embodied existence. It is for this reason that whenever we talk of disembodied survival, a person in this state would be

understood properly as a disembodied person, i.e. as one who originally had a body but now has none. But if our talk of person and same person in this state is to have anything more than a secondary sense, the belief in survival has to be wedded to an essentially incorporeal theory of persons. There are of course certain versions of survival which claim to free themselves from such a commitment, e.g. Hick's theory of instantaneous change of body at death,<sup>2</sup> and the rather disreputable theory that there may exist "nothing" - literally nothing - during the gap between death and resurrection.<sup>3</sup> And these, no less than the versions involving disembodied existence of the survivor at some stage or other, deserve our attention, in so far as they suggest, what we want to deny, that memory, in these circumstances at least, will be the sole (properly a sole) criterion of personal identity. In the space that follows, I will be concerned with examining what sort of possibility it will be if disembodied survival is possible at all. My second important concern, which will be considered in the section that follows the present, is to examine how far such possibilities can show that memory is the sole criterion of personal identity. In the interest of the latter, consideration of survival which, supposedly, do not involve disembodied existence will have to be deferred to that other section.

First, as to the possibility of disembodied survival. As I have just indicated above, a theory that purports to take this possibility seriously has two alternatives to choose: either to profess that a person is essentially incorporeal or to confess that our talk of 'disembodied persons' (and also of 'same person' in that context) can only have a secondary sense. The first choice will be difficult though congenial, the second choice will be relatively easier but embarrassing to the survivalists. Let us see how. The first choice would be difficult to establish. Criticism of this Cartesian way of thinking has been as old as the theory itself and several serious difficulties have lately been exposed by/too philosophers

numerous to mention. My plea for adding to these objections is to point out a limitation which has certainly captured the imaginations of the Cartesians but which, it seems to me, has affected the reasoning of some anti-Cartesians too. I shall point out this limitation by emphasising a connection, which in my view has hardly been understood in its right perspective, namely the connection between what is the case and what could be the case. (Although what I shall say about this connection may be true of varieties of other cases, I will confine my discussion to the cases of person and personal identity in so far as it relates to what is or could be a criterion of personal identity.) For I think that the connection between the two issues (I shall refer to them as the "actuality-question" and the "possibility-question" respectively) is more intimate, in fact of a conceptual nature, than has been supposed by both Cartesians and some anti-Cartesians alike. They have generally tended to treat this connection as one of absolute independence. In other words, they have supposed that there is no connection between what is the case and what could be the case, that the former has nothing to do with the latter and vice versa. The Cartesians, for example, have argued that although persons in fact are bodied beings this puts no limitation whatsoever on the supposition/they can be entirely disembodied (that is "non-embodied" in the way we have specified it). Some anti-Cartesians, on the other hand, have argued that even if persons could be completely disembodied, it does not follow that they are so.<sup>4</sup> The latter argument, although it has effect on the orthodox version of Cartesianism (that I am a 'thinking thing'), still leaves untouched the basic spirit of Cartesianism; for a moderate reconstruction of Cartesianism could insist on the possibility - not actuality- that persons are disembodied spirits. I suggest that the unsatisfactoriness of the above reply to Cartesianism is due to our supposition that the two issues of 'actuality' and

'possibility' are entirely independent of each other; and my further suggestion is that a more effective way of dealing with Cartesianism is by challenging this supposed independence which has supplied latitude to Cartesian and near-Cartesian theories. An initial first move in this direction may be to recall the obvious point that the denial of the "possibility-question" implies the denial of the corresponding "actuality-question" i.e. if persons could not be disembodied spirits, the assertion that they are would lead to a straightforward contradiction. (The connection between the two questions is not "independent", like, for example, that between leaves being colourless and gases being colourless: on the possibility score even if, say, it could not be that leaves are colourless it has nothing to do with gases being actually colourless.) I want to argue that the connection also holds in the opposite direction - though in a somewhat less stringent form, or less obviously so, at any rate. I.e. in the context of a particular type of thing, what is the case has often a lot to do with what could be the case. If what must be involved in considering the two issues of actuality and possibility is one type of thing or concept rather than several, then what that type of thing is has a considerable role in the issue of what that thing could be. In particular, what persons actually are has a lot to do with the concept of a person. To recall our earlier arguments (see i above) understanding or learning the meaning of "person" essentially involves knowing persons as they are (how else could we know them?) - knowing what sort of things they are. Only by knowing them as they are (being embodied as well as being thinking subjects), we learn the meaning of person and master the technique of applying it in the right contexts or to the right cases. Had we been taught to apply this concept to entirely different kinds of things, and not to ourselves - which is logically possible - then "person" would have had a different meaning than what it does have. This shows how what persons are is intimately connected with the meaning of "person". If it

were not for this important fact the problem of disembodied persons would not be a problem for philosophers. And it is for this reason that the idea that persons are incorporeal thinking things strikes us not only as counter-intuitive but also as conceptually odd. For it seems almost inevitable that any other use of this concept-i.e., if it applied to things and beings who are very different from ourselves - it would be a changed or extended use of the concept, whether or not it is too obvious to us. And the significant fact is that this extended use of the concept will have sense only because the original use has the sense it does, but not vice versa. Thus we are back again with our distinction between the primary and secondary use of our concepts. But the Cartesian or anyone who takes the idea of disembodied existence seriously, will not find this consequence very palatable. He will, on the contrary, profess that the talk of "person" in the two states (bodied as well as dis-embodied) has the same sense, - only that in the former state the idea of a body is added ab extra to the concept of a person. In effect he would be claiming that the disembodied state of the person is his "natural" state, and that though from time to time he may pass through several bodied existences the latter is not his "natural" state. (Significantly this is an exact antithesis of the Thomistic concept of a person or soul<sup>5</sup>). That this account of person will not be our normal concept of a person is clear enough. But further - more, if our analysis of the actuality - and possibility - questions holds, it would be difficult to see what meaning it may have. For even if it may have any meaning no one could be taught and so no one could learn the meaning of an incorporeal person. Since by supposition this is the primary sense of the concept of a person, a disembodied person ought to be independently identifiable

(since "person" is a substance-concept, and a substance must be so indentifiable) in order that the meaning of this person could be taught and learned - and this seems impossible. It is no good saying that the learning can be made possible by identifying and observing persons in their bodied state since, ex hypothesi, this state is not their natural state, and any learning through this state could only be that of a 'secondary' concept of a person and not of the primary concept. It follows therefore that an entirely incorporeal concept(non-embodied) of a person is unintelligible; and this is due to the Cartesians trying to dissociate completely the question of what persons could be from the question of what persons are, and unjustly disregarding the latter altogether while considering the former. Added to this difficulty, there is another crucial difficulty which an incorporealist has to face. If, at the risk of inconsistency, the Cartesian claims that the meaning of person can be learnt by observing and identifying persons in their bodied state the problem he has to face is: how could we know that there is a person (an incorporeal mind or soul) here and not something else, e.g. a shadowy, astral organism ? And even if this can be evaded, the more crucial problem still remains with us i.e. how can we know that there is only one person and not twenty or two hundred in a particular body having qualitatively indistinguishable experiences and mental states ? It is not clear how this question can be answered even in one's own case, not to mention the case of persons other than oneself. As Strawson expressed this difficulty: "How would each indignant soul, once this doubt has entered, persuade itself of its uniqueness <sup>6</sup> ? " The Cartesian has no answer to this: but there is a simple answer to be offered by those who emphasis the "actuality question" in the way I have, and the answer is: one body one person.

Here I must make a passing mention of a possible reply from

an unconvinced sympathiser of Cartesianism. He may argue that although we may have no ground for saying or knowing that there is only one person (incorporal) in a body it does not follow from this that there is not one person, but several, in that body. This is a profound objection, but there is to this a counter-objection which is equally profound and equally respectable. For it is not denied that there being only one person is logically possible; but what is important is a principle for applying this hypothesis - otherwise it would not be a usable but an empty hypothesis. For equally possible is the rival theory that there are many persons in the body and one can reasonably use this to make the counter-move in the Kantian Style: if you are allowed to invoke a hypothesis (the former) without feeling obliged to elucidate the principle of its application, nothing should prevent me from introducing the rival hypothesis (the latter), also unelucidated.<sup>7</sup> This line of approach may strike one as destructive and negativistic; but it has great relevance to us both of general interest and also of a particular nature. It is of general interest to our proposed thesis since its professed programme has been to consider what are, or rather what should be, the criteria of personal identity and not what personal identity is; it is also particularly significant to our present purpose, for since we are examining possibilities of what persons could be and how they could be thought to survive their death, only those possible hypotheses will be of any interest to us which can offer some principle of its application or which can be said to have some grounds at least for their assertibility.

Returning to the mainstream of our arguments, if an incorporeal view of persons as the primary concept cannot be consistently and intelligibly maintained, the Cartesian survivalist has got only the second choice we offered to him namely, that he has to confess that his talk of person and

same persons can have at best a secondary sense derived only from our actual concept (the bodied-subject-of-consciousness concept) which we have seen to be the primary concept. This choice may be embarrassing to him, but is the only way in which he can give a justifiable content to this theory. This choice not only offers him a good ground for presenting his thesis as an intelligible theory, it saves the theory from the charge of unintelligibility like what infects the <sup>Cartesian</sup> belief in what we described as "non-embodied existence". For, as we have argued, the only logical ground for saying that there could be disembodied persons is that we have the primary use of person (which is that of bodied subject-of-consciousness); but if this ground is ignored and the secondary concept of disembodied persons is given the status of primacy, then the latter would lose the appearance of sense\*. And once again it may be noted that what, in our view, gives this hypothesis the status of preferability (to that of non-embodied existence) is the fact that this hypothesis takes proper account of what persons actually are. If this analysis of disembodied survival is fair, and if the second choice is the only choice open to the survivalists, then the following consequences will follow: first, the possibility of disembodied survival will be an odd - conceptually odd, as we saw - sort of possibility since, in effect, it will be reduced to saying that person (and accordingly, same person) will not have the same meaning here as in the embodied state of existence. Secondly it will reinforce our earlier claim (which is the basic contention of the present thesis) that bodily continuity is the primary criterion of personal identity and thirdly, yet again

---

\* If a theory T is intelligible only in a secondary sense, and if anybody (e.g. the Cartesian survivalist) claims it to be the primary concept of a person, then in that sense T cannot be intelligible. (Whence it follows that persons cannot be said to be essentially non-physical)

memory can be seen to be a secondary criterion.

As regards the first of the above consequences, I claim to have given a fair amount of arguments in the present section. Let us try to see, by considering some representative theories of survival, that directly or indirectly the latter are committed to this consequence. I take it as an important, if not an essential, element involved in the belief in survival that what exists or survives after the death of a person is a soul or some such non-physical thing. In view of this, any such theory must be reasonably understood - unless a theory of this type qualifies itself otherwise\* - to believe in disembodied existence of the person involved at least for some time between going out of one body and entering into another. Quinton in his important paper "The Soul" can be fairly taken to be a reasoned representative of such theories, and since he does not make any of the described qualifications, he's committed to the belief in disembodied existence. i.e. he will be committed to saying that the soul not only goes out of and into bodies, it also continues in disembodied state during the transition. And like a consistent believer in disembodied survival, and quite to the consistency of this theory, he identifies his person with this soul. For his concluding remarks are not only that the soul is the 'essential constituent' of personal identity, "it is also what a person fundamentally is"<sup>8</sup>. And like we have seen (last section), in the interest of consistency and intelligibility, Quinton is obliged to accord the disembodied person a secondary status. In order to give plausibility to his belief in bodily transfer, he is found to confess that he is putting the concept of person (and I hope, personal identity) under "strains" of conceptual revision, or as he puts it himself, "extending" the concept. Whether he or anyone is justified in making this "extension", what this shows is that if in certain

---

\*Like supposing instantaneous change-of-body or claiming that if there is a gap between death and bodily resurrection, "nothing" need exist during that period.

circumstances the talk of person and same person becomes compelling, and if this talk has to have any intelligibility, this can only be possible by making the concept express a different sense from what it normally does. And admittedly this conceptual change does not and cannot make the new concept the primary concept nor can it show the incorporeal person as the fundamental concept since it was introduced (if not invented) by extending the normal bodily concept and as such could serve only as a secondary concept.

As another theory that allows the possibility of disembodied survival but only at the cost of a conceptual change with the consequent commitment to a secondary or dependent sense of these survivor persons, I choose the theory of Strawson in his Individuals. After outlining that excellent account that person is a primitive concept to which both corporeal predicates as well as mental predicates are ascribable, Strawson nonetheless felt obliged to do some justice to the conceivability of a "pure individual consciousness" that may continue to exist after the death of the person. Presumably since the concept of mind is a "derivative" concept, the continuance of a pure individual consciousness could not be granted the same logical status (like existing all by itself as an individual) as a person. Yet it is the notoriety of this concept that its continuance after the death of the body forces itself on the imagination as an irresistible logical possibility and since the distinctive - but, remember, not all essential -person<sup>al</sup>/faculties are attributed to this faculty of mind, the possibility of its continuance looks very much like the continuance of the person. So presumably, at least for the sake of the intelligibility and conceivability of the continuance of this 'pure consciousness', Strawson felt obliged to give an account of how this idea gets its intelligibility. And quite in keeping

with his theory (primitiveness of person), he rightly attributed the intelligibility of this hypothesis to there being the primitive concept of a person which is what we actually are, namely the "type of entity such that both predicates describing states of consciousness and predicates ascribing corporeal characteristics... are equally applicable to a single individual of that type". (Ibid pp 115-16 and 102). If this is so, and presumably since in the absence of the body the corporeal characteristics could not be truly ascribed, this post-mortem pure consciousness (call it disembodied person, if you like) can have only a secondary sense of being a person. For if it were not for the primary concept, as we have seen, this concept of a pure consciousness would not have come to have a use in the first place. Justly enough, Strawson calls this a secondary concept (Ibid p 102) and accords the individual consciousness a "logically secondary existence" (p 115). More significantly, he insists that such a concept could only be intelligibly formed from "within our actual conceptual scheme" (p 115 my emphasis), and this lends support to our claim that the idea of pure non-embodied existence, without any regard for what persons actually are, is unintelligible. This analysis of a pure individual consciousness makes no secret of the fact that if such a concept is to be called a person (al beit disembodied) at all it will be in a different sense of "person" from the sense in which actual, bodied persons are called persons. Presumably with this intention, Strawson refers to them as "former persons" (Ibid p 116). Strawson's allowance for disembodied persons and, particularly, his talk of them as persons has been the subject of serious controversies lately, and it has been generally regarded as an obvious inconsistency in his own theory of persons.<sup>9</sup> Flew, for example, is not willing to allow Strawson the "easy imaginings to assume that

his putative disembodied beings would be persons ..." (op cit p 232, my emphasis). It seems to me however that, with the exception of Cowley's criticism - which purports to rule out the logical possibility of a dualist's mind (conceived as an entity) being derived from the concept of a person - these criticisms are based on the assumption that Strawson used "person" in the same sense in both the cases of disembodied persons and of the normal bodied persons. And if he did so he could not escape the force of those objections. But as my reading of Strawson's position shows, there is for him a definite way of escape. For on my showing, he used, or at any rate can be said to have used, the concept in two different senses in the two contexts - that whereas it has its primary use in the bodily context, the concept has only a secondary, dependent use in the disembodied context. Further more, it is significant to note that although he does indeed speak of disembodied person for a couple of times in the 3rd chapter of the Individuals (p 103), Strawson explores the possibility of disembodied survival by introducing the concept of a "pure individual consciousness" and keeps on using this concept more often than the concept of (disembodied) person; and the few times he used the latter to refer to the survivor, he qualified this noun with the adjectives "former" and "disembodied" with a purported emphasis on the "dis". Happily, however, in his latest contribution in this field of the literature he has apparently given up the idea of disembodied person altogether.<sup>10</sup>

Another theory of disembodied survival, worth considering here, is that of Aquinas. As I see it, Aquinas can be fairly described as the ancient precursor to the Strawsonian position - both with regard to disembodied survival and with regard to the concept of a person. Like Strawson, or rather unlike the Cartesian dualists, Aquinas maintained that person is a psychophysical unity, one

single substance and not composed of two individual substances like the mind and the body. And like Strawson again, although he allowed for the possibility of disembodied survival of the pure ego (the soul), he nonetheless was committed to accord the latter at best a secondary, dependent status; for both maintained that the requisite "individuality" of the (disembodied) soul or ego is due to the fact that it retains this individuality (or the natural orientation to "inform" the body - Aquinas) from having been a (an embodied) person.<sup>11</sup> (Aquinas even went on to make the soul dispositionally dependent on the body in the sense that although the soul in its disembodied state was capable of performing its higher, sensitive activities, it could not exercise this capability without its union with the body, and so is in need of this union.<sup>12</sup> But while Strawson could not, at places, resist speaking of this survivor soul as a person, Aquinas, more carefully and consistently, had no hesitation, in declaring that the soul, in the state of separation from the body, can't be a person.<sup>13</sup> Nor, according to Aquinas, can this disembodied soul be even a distinctive, feature of personhood, since this the soul can be only in its natural state or in its state of "perfection" which, according to Aquinas, is its state as the "form" of the body, and this it can be only in its union with the body. It is for this reason that Aquinas holds that the disembodied state of the soul is not its natural state (suma Theologica Q79A2); it is what he calls praeter naturam or the state beyond nature. All this confirms my contention that the idea of disembodied persons, if intelligible, is so only in a secondary sense and that the concept of a person, as applied in this case, is applied in a substantially different sense. If our arguments in this section hold, then it cannot be claimed that our use of the concept has the same sense in this case as it has in the normal case of bodied

existence. To believe this will be to forget the fundamental distinction between the described primary sense and the secondary sense of the concept. Only with this prerequisite can a hypothesis of disembodied survival be intelligible and the intelligibility of any such hypothesis must owe allegiance to the requirement of there being the primary use of the concept in the normal context, for otherwise we would be losing the very ground for stating the supposed hypothesis as an intelligible theory. The truth of this will be confirmed further if we consider some of the arguments usually given for disembodied survival.

There are two essential requirements to be considered with regard to the disembodied survivor. First, he should satisfy certain minimum conceptual requirements of being a person and second, since he is a disembodied person, enough content should be given for his identity through time. The second question, which concerns the criteria of personal identity( as applied to the disembodied case) will be considered in the next section. I shall now confine myself to the first question. Admittedly, not all the predicates that apply to persons could be ascribed to a disembodied person for the obvious reason that he does not have a body now. What is important, therefore, is whether enough could still be ascribed. It seems perfectly conceivable - especially if one imagines the disembodied person to be oneself - that most of the P-predicates of Strawson can still be intelligibly applied; it is also argued that even some of the corporeal characteristics can still be applied. It will be recalled that it is not our intention to deny that these are logically possible. But what we will point out is that the application of most of these predicates or characteristics will inevitably involve certain complications. And this the advocate of disembodied survival can't deny. For example, it is said, without a body a disembodied person could still be able to move himself from place to place;

but how can he do this. Not in the way the normal persons do this i.e. by moving their bodies; the only way in which a disembodied person can do this, it is said, is by "just trying to do so". Similarly he can lift a table for example by "just willing to do so". Harrison could imagine himself lifting and moving the limbs of other embodied persons by simply exercising his will.<sup>14</sup> Now without needing to enter into conceptual complications, it may be simply said that by this means we are endowing the disembodied person with certain special, perhaps magical powers. For will it not be a magical power if an embodied person could do things in the described way, If this is so, it is easy to see what sort of persons these disembodied persons would be; and it becomes preposterous to deny that we are not making a change, or at least an extension, in the concept of a person. Let us anticipate a possible objection to this only to get it out of our way. It may be said that if some persons, normal persons, do possess certain special powers like this (e.g. a group of magicians or black magicians perhaps), won't we still call them persons and will the concept have a different sense in their case? Perhaps we will not say that, but if the entire mankind were endowed with such extraordinary powers(perhaps they would not be called "extraordinary" then), then even if they were called persons, <sup>been</sup> that concept would not have / the same as the concept is today. Perhaps, Strawson would have to stipulate a third category of predicates in his analysis of person. Or, imagine all over the world from now on every- body is endowed with such extraordinary powers, e.g. everybody could see anything happening anywhere, bring out changes in any environment and the like. In such case, persons would be very little different from gods. Any if there is a difference between the concept of a person and the concept of god, what would we say except that the

concept of persons has to be revised and at least extended in its application to us in this imagined state of affairs ? The case of the disembodied persons being endowed with certain special powers is analogous to each of our imagined cases. And thus inevitably the possibility of disembodied persons is bound to effect conceptual change. Furthermore, as we have seen, the concept of person when applied to these cases will have a different, secondary sense. This is evident from the admittedly essential requirement that for many of the predicates to be applied to them, they must have been acquired or "inherited" from the embodied days. Since it is the possession and / or exercise of these powers that makes them persons (albeit disembodied person) and since for this they have thus to depend on their prior embodied existence, it follows once again that they will have to be persons(i) in a different sense and (ii) in a dependent sense - despite the illusory confidence of their authors that they are persons in the same sense as embodied persons are. It follows therefore that our survivors cannot have the same logical status as we enjoy (consequence of (i) above) . nor, to the disappointment of the Cartesian survivalists, can "disembodied person" be the primary concept of a person so that the nature of persons could be essentially incorporeal ( consequence of (ii) above).

Section (iv): More About Disembodied Persons.

In the last section, we distinguished two supposed theses, the thesis that the only persons that there could be are disembodied persons (this was described as the belief in "non-embodied existence") and the thesis that there could be disembodied persons surviving their bodily death (described as "disembodied survival" or simply "survival"). We argued that the first thesis was unintelligible since although it owed its apparent intelligibility to our talk of person and same person (in a secondary sense) in certain circumstances, it deludes itself to the status of primacy - presumably by taking the fact in its face value that we use the words "person" (and "same person") in these cases as well as in the normal cases. Further, we argued that the second thesis - "disembodied survival" - is an intelligible hypothesis, but only so in a secondary, dependent sense, since if we did not have the primary, bodily concept of a person, we could not give any sense to the idea of a disembodied person. Now, taking this thesis to be intelligible in the way it is, we must consider how it relates to our account of personal identity. As we have indicated earlier, the possibility of disembodied persons and disembodied survival appears to collide head on with our stated claim that memory is only a secondary (dependent) criterion of personal identity, and that even its being the sole criterion (a sole, to be more exact) in certain circumstances is highly doubtful. For if there has to be such persons at all, the criterion of identity between a later (disembodied) person and an earlier (disembodied) person and that between a later (disembodied) person and an earlier (embodied) person will have to be memory and memory alone - since, obviously, there is no body in the former case and no bodily continuity in the latter.

Now, if we are right in saying that the idea of disembodied persons is intelligible only in a secondary sense and that the concept of a person as applied to this case has only a secondary sense, then it seems

perfectly reasonable to say that memory, in so far as it serves as the (only) criterion for the identity of these persons, is only a secondary criterion after all. And if from the fact that there could be disembodied persons it does not follow that there could be pure "non-embodied persons", then whatever criterial role memory might have in the former case this cannot be its primary criterial role in determining personal identity, and moreover, since the idea of pure "non-embodied existence" has been seen to have no clear sense, the consequent idea of memory being the sole criterion or even the primary criterion of personal identity - in so far as it draws force from this idea - can have no clear sense either. However, let us consider the case of disembodied survival in its own merit, and see if the result can be otherwise with regard to the role of memory.

Obviously, if we are to suppose that a person could continue to exist after his bodily death, we have to give sufficient content to the idea that the "survivor" is the same person as the pre-mortem person he claims to be (or is claimed to be). We have also to give content to the idea that he at any particular point of time is the same (disembodied) person as any earlier (disembodied) person. For the idea of a person, being that of a continuant, carries with it the notion of reidentification through time. Clearly, the only ground for saying that the survivor is the same person as any earlier-pre-mortem or post-mortem - person is by saying that he ~~remembers~~ doing actions and having the experiences of that earlier person. Let us grant that such a person can remember and (which is more important) say that he remembers, i.e. can make memory claims. In granting this, we will ignore certain obvious "difficulties of how he can make these claims or utter the words "I remember ... " without having the appropriate bodily organs.

It should not be objected here that even if the survivor cannot say

"I remember ... " it does not follow that he does not remember the actions and experiences of the earlier person. For it will be recalled that here we are not concerned with whether the survivor is the person as the earlier person but with what criterion can we have for saying that he is, and whether, as is generally believed, memory is that criterion. Further when our concern is basically what criterion can be there for the identity of these persons, we are bound to consider this possibility in the third person, for even though the case of disembodied survival seems much stronger when considered or contemplated in the first person - indeed it is this way of looking at the matter that gives a basic source of intelligibility to the supposed thesis - yet the question of criteria is irrelevant in this case, for as we have said before and as will be generally agreed,<sup>1</sup> neither the knowledge of self-identity (one's own identity) has to be based on criteria nor do we use any criteria in making first person memory-judgments which imply that identity. It is imperative, therefore, that any discussion of the criteria of personal identity (and so of personal identity, since the former consideration is indispensable to it) must include a closer analysis of what is true of others when they are said to be identical with any earlier or later persons. And assuming, as we do, that bodily continuity and memory are the only two fundamental criteria of personal identity, the consideration of what is, or that anything is, the criterion of disembodied persons requires that they be able to make memory claims. In a solipsistic world of the disembodied person,<sup>2</sup> where no other disembodied person can be heard or listened to, and indeed nothing can be known as to whether they can remember or even whether they are there at all, the suggestion that memory or anything is the criterion of identity has no clear sense.

So, let us ignore the difficulties and grant that a disembodied person can make memory claims in whatever attenuated sense the authors of this hypothesis may content themselves with. I shall, however, main-

tain that this will hardly show that our "survivor" remembers what he may thus claim to remember, . . . much less that memory is an independent criterion of personal identity. For we have already argued,<sup>3</sup> the use of memory as a criterion of personal identity presupposes (and so depends upon) the fact of bodily continuity. This dependence, it has been made clear, is explained not by saying that whenever we make personal identity judgments on the basis of memory we do so by checking the genuineness of the latter with reference to bodily continuity, but by saying that in particular doubtful cases the possibility of this checking must be available to us. Moreover since the distinction between ostensive memory and real memory is a conceptual matter, and since memory being a criterion counts on the possibility and probability of any of the former being a case of the latter,\* the described dependence on bodily continuity is an essential element in the claim that memory is a criterion at all of personal identity. And it will soon be clear that this fact of the described dependence stands out as an additional destructive factor in the case of disembodied persons. For in the absence of bodily continuity, the possibility of checking on the genuineness of memory and also of the correct use of memory-words (for this also requires the continued presence of the body, see pp 161-2 above) would be ruled out, and yet the possibility of a disembodied person's being mistaken in these respects cannot be ruled out, (unless, of course, we are prepared to stipulate that a disembodied person simply can't be thus mistaken!) Further, as we noted above, it is the availability of bodily tests which, in the face of reasonable doubt, would help to set the doubt to rest - and which

---

\*I.e. what makes any particular memory claim a justificatory evidence of personal identity is the presumption that this may be a case of real remembering. I have argued that memory in this weak sense should be understood to be a criterion of personal identity if no question is to be begged. But if this is so, and if what has just been said in the text about knowing an ostensive memory to be real memory is correct, then this criterion's implicit dependence on the bodily criterion is far too obvious.

thus underwrites, as it were, the role of memory as a criterion of personal identity; but with the lack of availability of this test in the disembodied case, memory's role as a criterion would be extremely suspect. It is therefore only to be expected that even some devout preponents of disembodied survival (and also some of these who would subscribe to it in some way or other) should have been obliged to concede that memory could not be a criterion of identity for disembodied persons or that, even if we may talk of there being the "same person" in many such alleged cases, there are nonetheless no criteria of identity here on which this talk is based. As a clear example of the former I take Harrison (in his paper on "Embodiment of Mind", PAS 1974), and as a probable case of the latter one can cite Malcolm (in his unpublished paper on Cartesianism referred to before). For largely because of the reasons we have given above for the dependence of the memory criterion on the bodily criterion Harrison "agreed" that "there must be something other than memory which determined whether I was identical with the person I seem to remember I once was,"<sup>4</sup> and even went on to imply that memory was an "impossible criterion" of personal identity.<sup>5</sup> And if he is right, there may be some other which, obviously, is not bodily) criterion of identity for the disembodied persons; but what interests us for the moment is that memory is not that criterion, since our aim here is to make an even weaker claim that the possibility/intelligibility of disembodied survival does not show, as is usually believed, that memory is the sole criterion of personal identity. Malcolm, on his part, appears to believe that our talk of a "person" and of his being "the same person" as the earlier, Robinson, whom we have known to have been dead and buried, is intelligible only in a secondary sense, and that, part of the reason for this is that, in the absence of the body and bodily continuity, there is no criterion of personal identity here.<sup>6</sup> Although I do not incline to share unambiguously the view that memory is an "impossible"

criterion of personal identity or that we have no criterion in the disembodied case, I can certainly see the basic force of one's arguing to this effect in the disembodied case. For if memory is a dependent criterion and if what it depends on is dropped out in principle, then inevitably it will look pedantic not to say that the "memory" cannot be used as a criterion in this case; and equally, assuming, as we do, that bodily continuity and memory are the only two criteria\* that can be used to determine personal identity, it would look compelling to say that we use no criteria in this (disembodied) case. And I would largely agree with this conclusion, especially since it is in effect my basic claim that we will have no justification to suppose that there are disembodied persons and that each of them is the same person as some earlier embodied person. But this must not be taken to imply - as it might be<sup>7</sup> - that the idea of persistence through time of disembodied beings is unintelligible and logically impossible. However, I have already explained (last section) what sort of "possibility" it will be and what sort of intelligibility it will have (and our way of explaining the notion of disembodied survival can now be seen to have yet another ground, namely that our normal criteria of personal identity cannot be applied here), and accordingly I do not see any prima facie difficulty in saying that we do have some criterion of identity for disembodied persons - but only in an equally attenuated and secondary sense. In a sense the specification of some

---

\*The possible suggestion that there could be a non-physical causal connection to explain the identity of persons (cf Harrison, op cit p 49) will not detain us here, since all we are interested in is that the notion of disembodied survival can't show that memory is the sole criterion, and that is conceded by this suggestion. Besides, if there were such a connection (and I take it that this would be the criterion in the bodily context as well as in the disembodied), it would, like the soul, be something unobservable and private, and so cannot be used as a criterion; it might (which we do not deny straightaway) contribute to the view that personal identity is something unanalysable, but that certainly is a different matter.

criterion seems imperative, for although there being disembodied survival (i.e. the idea being possible or intelligible) does not require that there should be some criteria for identity, yet saying that there are (i.e. that a supposed disembodied survivor is the same person as an earlier embodied person) - as Malcolm agreed that we do say this - certainly does. So I suggest (because I am also prepared to believe that in certain strange but compelling situations we do say this) that the criterion for saying that a later disembodied person is the same person as an earlier bodied person is memory - but with an important difference. For although, like the normal cases of embodied persons, we will rely here on the memory claims, yet, unlike those cases, we could not possibly check upon the genuineness of any of these claims. Thus, in the disembodied case, we could never know if ever a memory claim was true, consequently we could never know if any disembodied person was identical with any earlier person or with the pre-mortem person he may claim to be. We have said that it is the fact of there being such things as real memories that makes memory (memory claim or ostensive memory) a criterion of personal identity, indeed it is not just this fact, but the fact that there are known (at least, can be known to be) such memories that makes memory a criterion at all. This condition is easily satisfied in the case of normal bodied persons; but in the case of disembodied persons this condition could never be satisfied. It would be no longer any good saying (or relying on the fact) that people's memory claims are generally true; for the fact of these claims being generally true is not anything in the nature of these claims themselves (or in the utterance of the words "I remember ...") but in the nature of who makes these claims. And once there is a radical change - as there obviously is in the disembodied case - in the nature of who makes the claims, that rule (of general reliability) will not do, or, at least, it will come under fresh review (for more about this see pp.249-51 below). But no way of any such

review or test is open to us. Nor will it do to say that, since the disembodied persons are each of them the same persons as some embodied person who have only suffered bodily death, the general reliability of their memory claims should not be suspect, for presumably, not only the learning of the memory-words and their correct use but also the general ability to make "true" memory-claims will have been "inherited" from their former state of embodied existence.<sup>8</sup> This will not do, since what is at stake is whether at all any of them is the same person as any pre-mortem person. It follows, therefore, that in the context of disembodied survival, the point of saying that memory is a criterion of personal identity will have lost its original purport and force. However, if nonetheless the idea of disembodied survival is intelligible and if, therefore, it is intelligible to say that a particular disembodied person (e.g. Malcolm's disembodied Robinson) is the same person as a previous embodied person, then the only point of saying this would be that he "remembers" (makes memory claims, or what looks like memory claims) what that earlier person had done or felt. My suggestion is that since this is what we would do, or would be inclined to do anyway, in certain strange but compelling circumstances, and since we would be doing this on the basis of what may look like memory claims, it will look pedantic not to say that we are using a criterion and that this criterion is memory. But I shall also claim that the "memory" that we thus use as our criterion for identity of disembodied persons is only a pseudo-criterion with no possibility of its ever being able to play the proper criterial role. For as a criterion of personal identity, we have just underlined the vital difference in the role of memory in the disembodied context. Indeed, if our above analysis is correct, its role as a criterion of personal identity here is only a degenerated one - "degenerated" because it does not have its usual criterial function, but only the appearance of one. For a criterion is one that justifies us in saying that something (of which

it is supposed to be a criterion) is the case; but "memory" in the disembodied case, though it looks like such a justification and inclines us to say that a supposed disembodied person is the same person as an earlier embodied (or disembodied) person, can hardly be said to be a justification in the strict sense - for the simple reason that no ostensive memory in this case can ever be known to be real remembering, and so no disembodied person can ever be known to be the same person as an earlier embodied person (or, indeed, as any earlier person - embodied or disembodied). The conclusion to which this leads us is that the memory criterion in the disembodied case has come under the strains of a double disadvantage: earlier, its only weakness was seen to be that it is a dependent criterion - dependent, as it was, on the bodily criterion - and it is in this sense that it has been usually called a "secondary" criterion, but now, in the context of disembodied survival, the secondary status which we were willing to accord to it makes it appear much worse; for what we would be counting on, if we did count on "memory" as the criterion in this case, is not "memory" in its usual sense, nor even a memory claim,\* but something that only looks like memory claims with no possibility whatsoever of these claims being known to be real remembering. Interestingly, what these memory-like claims lead us to - if they do lead us to anything - is a similar world of appearance, with no possibility of its ever being known to be real; and the supposed world of disembodied survivors is nothing but such a world. Clearly, the intelligibility of such a world cannot be denied nor can it be denied that there may be such a world: but what is not clear is whether by believing in such a world we are not inventing what is not the case. However, what follows from the

---

\*For in the normal case, a memory claim is taken justifiably to be a reliable guide to personal identity, which the claim in the present case is not.

foregoing discussions is this: that the possibility of disembodied survival cannot show that memory is the sole criterion of personal identity - certainly not in the sense of being the only criterion that is used and can be used (our sense 2 above) for the purpose of reidentifying persons, and even if it may seem to be the only criterion in certain cases (our sense 1 of 'being the sole criterion') viz, in the case of disembodied persons, it would do so only as a degenerate criterion. For, a closer analysis of the situation reveals that, except for taking the memory criterion in the last sense, there would be no point in saying that memory is a criterion at all in determining the identity of the disembodied persons. Having achieved this so far, let us now examine a slightly different area of "survival" the possibility of which may be yet another source of the idea that memory is the sole criterion of personal identity.

As we noted in the last section this area supposedly includes the possibility of survival without any commitment to the disembodied existence of the alleged survivor at any stage. The hypotheses that purport to give content to this sort of possibility are (1) that of instantaneous change of bodies by the person and (2) the hypothesis that between 'death' and 'resurrection' or 'reincarnation' - even if a long temporal gap is involved - there is to exist "nothing". In both these hypotheses, if anything is to count as the criterion of personal identity, it is to be memory since apparently the bodily continuity is disturbed by the fact of death (bodily death) of the person who immediately or later on is to come to exist in a different body. I shall assume that it will be a different body despite the general belief (e.g. of the Christian resurrectionists) and even the insistence (e.g. St Thomas Aquinas<sup>9</sup>) that the pre-mortem body and the post-mortem body are the same; for I can see no good reason to say that the resurrection body is the same body rather than a different

one or even an exactly similar one. Furthermore I shall so assume - and it won't be unnatural to so assume - since some major religious traditions, like Hinduism, explicitly believe in personal survival (reincarnation or rebirth) in numerically different bodies, and even according to ancient Buddhism these subsequent bodies need not be human bodies but may be those of animals or birds etc. (I suspect that this belief in metempsychosis is not limited to this religious tradition alone.\*) Added to this, the hypothesis of instantaneous change of bodies, as it was proposed by Hick, assumes the two bodies not only to be numerically different ones but also to be existing in numerically different spaces altogether which may not be spatially related to each other.<sup>10</sup> Thus, with the assumption of different bodies being made explicit, the aforesaid hypotheses (1) and (2) will point to the memory criterion as being the sole criterion of personal identity. Of these, (2) is only the general theory of resurrection and reincarnation with the belief in disembodied existence of the resurrectee during the transition dropped out. But if, as (2) expressly assumes, there is to be nothing existing during the transitional period, and so the person simply goes out of existence at 'death' and again comes into existence at 'resurrection', then clearly it will not be a case of personal identity in the first place. As Locke had said, "... that which had a different beginning in time and place ... is not the same, but diverse".<sup>11</sup> And nevertheless the fundamental problem of how to check the veridicality or reality of the memory claims of the alleged survivor would still remain - if our arguments above are right. For with the original person and the subse-

---

\*Obviously the idea of metempsychosis would involve greater complications and implausibilities; but we need not go into these, for it will suffice to point out that whatever difficulties will be found to infect the idea of resurrection (in human bodies) will also evidently affect the idea of metempsychosis and such like.

quent person having different bodies, there will be no possibility of checking\* this. Therefore, whatever intelligibility and intuitive appeal this theory may have, we will have no grounds whatsoever for saying or supposing that the "survivor" is a survivor - that he is the same person as the pre-mortem person in question. But then if a criterion is a justificatory evidence for saying that something is the case, it follows that the supposed possibility of bodily resurrection and reincarnation - even in the qualified form as (2) above - will not show memory to be the criterion of personal identity and that it will even undermine its claim even as a criterion. Indeed, as has been shown in the case of disembodied survival, (similar reasoning will show that) our belief in, and talk of this kind of survival as (2) too will at best show that memory is a degenerated criterion. Now since this consequence follows from the lack of possibility of checking particular memory claims in the absence of bodily continuity, some might try to avoid this difficulty by suggesting either that there is a possibility of checking the memory claims of a survivor or that no checking is necessary, the latter suggestion drawing support from the fact that people's memory claims are generally true. Both these suggestions have in fact been made explicitly by G.C.Nayak<sup>12</sup> in support of disembodied survival and bodily reincarnation respectively. I shall show, by considering them, in the way they have been made and (by implication) also in the way they could be made, that these suggestions will not do the trick, since while the former is irrelevant the latter is a gross over-simplification. Nayak makes the first suggestion in order to show how a memory-claim

---

\*Nor will it do to say that a memory claim of such a survivor can be checked by its being found to be true that someone had done what is thus remembered; for the real test of x's memory claim is not simply that someone did what x claims to remember but that he did it. And this is not possible in the described context.

about the supposed former life of a disembodied person could be checked (and I also take it to be the procedure by which the memory-claims of an embodied survivor might be thought to be checked); and he appeals to a kind of "verification procedure"<sup>13</sup> which the person himself may take resort to. It is significant that the supposed procedure has been conceived in the first person (and, as will be evident, in the case of disembodied persons this procedure can only be conceived in the first person). The procedure is this: I may "remember" that in my previous embodied state I kept some valuables in a secret place and I may verify this by visiting the place and finding the treasure. Thus, it is argued, I shall be "convinced" - presumably, of my identity with the previous person. This may be true, but very uninterestingly so. For what we are concerned with is what the criterion/criteria of personal identity is, and it has been seen before that first person judgments implying self-identity can't show us any. The really interesting and relevant question is: whether I would be "convinced" that someone else is the same person as a previous embodied person if he carried out this verification procedure. But this question can have no definite answer from Nayak's suggested procedure. For admittedly, I would not know if he (another disembodied person) carried out this procedure. The procedure may, of course, seem to have some plausibility if it were carried on by an embodied survivor (though Nayak does not use his argument in this way); but that will not show that the survivor is really remembering and, so that he is the same person as the previous person. However, more about the last objection will be brought out by examining Nayak's second suggestion, namely that no checking is necessary for any of the survivor's memory claims. (Nayak is there concerned with the embodied survivor who claims to be a reincarnation of an earlier embodied person). For, he argues, there are "good reasons for believing the memory claims that are made with sincerity and conviction to be veridical more often than not,"<sup>14</sup> and he "fail(s) to

see why this should not be true also of those few memory claims of earlier lives that are made with sincerity and conviction".<sup>14</sup> I submit that whereas Nayak is certainly right about the first part of this claim, he is guilty of oversimplifying the matter with regard to the second part. To see this, let us understand carefully what the first part of the claim amounts to. Surely, there are good reasons for believing that sincere and confident memory-claims are generally true. Indeed, as we had seen, this can be taken even to be a necessary truth. Shoemaker who made this claim explicitly, reasoned that this is one of those "general facts of nature" which must be assumed in order for our concepts to have significance.<sup>15</sup> In particular, if it were not a "general fact of nature" that sincere and confident memory claims are generally true, no one could possibly make and none could be understood to make memory statements at all. Now, given this 'general fact', a sincere memory claim of any particular person is not to be doubted unless there was reasonable ground for such doubt and checking. But certainly, this is no reason for saying that the sincere and confident memory-claims of his past life by an alleged reincarnate is not to be doubted and not to be checked. For firstly, we must remember that the truth of the sincere and confident memory-claims is a general fact of nature, and secondly, this general fact is compatible with there being reasonable grounds for doubting any particular memory claim. And it will hardly be denied that the case of an alleged reincarnated survivor (or a "resurrectee") recounting events and actions of "his" past life is not a general fact of nature. How, then, can a general fact explain or account for a phenomenon that is not so general and admittedly "few"? Further, I believe that the "general facts of nature" which Wittgenstein<sup>16</sup> refers to, and which subsequently Shoemaker appeals to, are closely connected with what may be called a semantic feature of our language. This can be explained by saying that, for example, in order that certain statements may be made and understood

as memory-statements, certain semantic conditions must be satisfied, viz, that the words "I remember ..." must be uttered by persons and that the utterance of these words must be correlated with certain happenings in their pasts and so on. And what is, perhaps, equally important is the fact that these correlations must be known to hold in most cases, for otherwise, not only the use of memory-language could not be taught and learned, but the general reliability of the memory claims can never be guaranteed. Now, in the case of the alleged resurrectees of reincarnated persons, what are the semantic conditions that might give content to their "remembering" their earlier lives? In the absence ex hypothesi of bodily continuity, none of the "remembered" actions of a previous life done by any previous person in that life (if such could be known) could be reasonably said to be happenings in their pasts - much less, known to be so. It follows, therefore, that not only are the cases of resurrection and reincarnation not general facts of Nature, they also supply a reasonable ground for doubting the alleged memory claims in these cases and call for the need to check at least some of these claims. Besides, this reasonable ground becomes a strong ground in view of the fact that such cases are not only admittedly "few" but are certainly abnormal. And as Wittgenstein said "It is only in normal cases that the use of a word is clearly prescribed; we know, we are in no doubt, what to say in this or that case. The more abnormal the case the more doubtful it becomes what we are to say."<sup>17</sup> It follows, therefore, that in the cases in question checking of memory claims is needed and that the normal rule of general reliability will not apply to the memory-claims about some supposed "past lives" (Nayak, in so far as he argued to the contrary in his second suggestion, was making the mistake of treating the normal case of our actual life on a par with the evidently abnormal case of reincarnation). But since, in the absence of bodily continuity, there is no possibility of such checks, there would, in such cases, be

no ground for saying that a person is really remembering things and events in his past life rather than showing an excellent feat of retrocognitive clairvoyance. Penelhum has argued that it would only be a matter of option as to what to say in these cases and that the identification simply on the basis of these memory-like claims does not have to be made,<sup>18</sup> to which we add that it will be unreasonable to make such identification. Incidentally, the proposed "verification-procedure" discussed above may seem to be more plausible if known to be performed by an embodied survivor. A Charles in the 20th century may claim to remember what a Guy Fawkes in the 16th century had hidden in a secret place, and may also be able to visit the place and find the treasure. Yet this will not show that Charles really remembered. For what is required is not simply that what he "remembered" be found to be true, but (since it is a case of personal memory) also that it is he who had hidden the thing. In the unavailability of bodily continuity, all necessity for saying this is lost and no justification for saying this is forthcoming. If this is so, and if our analysis of the criteria role of memory is sound, then it follows once again that the alleged cases of resurrection and reincarnation, being themselves dubious cases of personal identity, fail to show that memory is the sole criterion of personal identity in any sense.

Now, a close look at hypotheses (1) and (2), and how the foregoing has revealed them. It has now become pretty clear that nothing really has been gained by dropping the idea of the intermediate disembodied existence (2) nor, indeed, by stipulating the idea of instantaneous change of bodies (1). For admittedly, both in the theory of instantaneous change and in the theory that "nothing" exists between death and resurrection, different bodies are involved. And inevitably, despite their obvious advantages concerning identification (individuation) of persons and their memory-claims, they would still fail to supply a

definite answer to the fundamental question of whether they really remember and so are the same persons as the pre-mortem persons. We have just said that, in the absence of bodily continuity, the answer to this question can at best be given in either way. And so the possibility of such survival would not show memory to be the sole (or, as we have argued, even a sole) criterion of personal identity: by a similar line of reasoning as the above, what it may show is either that memory is not a criterion in these type of cases or that, if it is, it is a criterion in a degenerated sense. And here again, since the plausibility of these hypotheses depend, as it does, on this degenerate criterion, the talk of "same person" and even of "person" in these cases will have suffered similar degeneration. This is more particularly true of the second hypothesis referred to above. For by making the plea that there may exist "nothing" during the gap between the pre-mortem and post-mortem existence, it not only suggests a different account of the concept of personal identity, it also appeals in effect to an entirely different concept of a person - as that of a gap-inclusive entity - something like television serials or orchestras, and the like, which exist by instalments,<sup>19</sup> as it were. This obviously, will not be the normal concept of a person which is the concept of a continuant, since this will be the concept of an entity which would be said to be continuing (identical with its(?) earlier "instalments") without being continuous. No doubt, therefore, this will mean a major conceptual change or revision; though, it will be argued in the next section, such change is neither justified nor necessary. Further, since, in actual life, persons often do suffer from loss of memories, it should also be possible that in each (or at least in some of these) different "instalments" (person-stages, as they can also be described) there might occur such loss of memory. In other words, one 'instalment' or person-stage might not contain any memory of its(?) previous instalments. In the normal case no such possibility

of memory-loss obstructs reidentification of persons; for there is, in that case, the other criterion of bodily continuity to ensure who someone is. But with the essential absence of this, in the supposed cases, the described loss of memory would make identification impossible; and if it is impossible to say who  $P_2$  is, it is pointless to suppose that he may be the same person as  $P_1$  rather than  $*P_1$ .<sup>20</sup> In other words, since the gap-inclusive concept of a person must make it intelligible for the described loss of memory to be possible, and yet subscribe to the view that there is no other criterion except memory, it would be committed to saying that there may, in these cases, eventually be no criterion of personal identity.\* And a theory which, thus, in effect amounts to saying that there may be personal survival without there being any criteria of identity, can only give us not only a secondary use of the concept of personal identity, but also a degenerate concept (in the same sense in which memory, in these cases, has been shown to be a degenerate criterion). In view of the basic agreement in their claims and contents, the last conclusion will apply as well to hypothesis (1) as it does to (2). And if the theory that memory is the sole criterion of personal identity leans, as it often does, on survival theories of these types, so much the worse for that theory.

---

\*Moreover, even when a subsequent instalment-person "remembers" what an earlier instalment-person did, this will not show that he is really remembering; for, as is possible in actual cases, he might be q-remembering the activities of an entirely different instalment. And since the only remedy for such eventualities is the possibility, at least, of an independent check of the memory-claims, the unavailability of this minimum requirement will show neither the fact of personal identity nor that anything - let alone memory - is a criterion.

Chapter 5

CONCLUDING ESTIMATE

Section (i): Conflict of Criteria  
And the Relevance of Puzzle Cases

So far, in the preceding chapters, we have analysed the nature of the problem of personal identity and made a comprehensive survey of the problem as it has been understood and dealt with by various traditional as well as contemporary philosophers. In the course of this enterprise, two major claims of consequence have emerged and have been consistently maintained. The first is that the problem of personal identity is not, as it has appeared to most traditional philosophers, a problem of defining personal identity, but is one of criteria. Once the right criteria are specified for making identity judgments about persons, we will have done all that there is to it; and the question of what is personal identity and how it can be defined, being itself unanswerable, should better be left unanswered, for it has been shown that no non-trivial definition of personal identity is possible and hence that no definition of it would be philosophically useful. In view of this, most traditional theories (and even some contemporary ones) have been seen to be fundamentally mistaken, since they were trying to define personal identity in some way or other while what, indeed, they were giving, in the guise of definitions, were only the criteria for saying that a later person is the same person as an earlier person. (It has also been noticed that although some traditional philosophers, e.g. Butler and Reid, rightly realised that personal identity cannot be defined in terms of what is only a criterial evidence thereof, they were still mistaken in so far as they reasoned, on that ground, that consideration of such evidence, viz, memory, is unimportant or worthless, see Ch 2 sec ii.)

Our second claim has been that, of the two main criteria of personal identity, the bodily continuity criterion is the more fundamental one and that the memory criterion is only a secondary criterion in as much as it has to depend on the former in order that it may be used as a criterion at all.

In view of our above two claims, and particularly of the second, it can be seen that the possibility or the fact of conflict between the two criteria should not be worrying as it generally has appeared to be. For, on our analysis, any such conflict will either decide the matter in favour of the bodily criterion or would make the question of identity - if not a matter of arbitrary decision - amenable to a secondary and somewhat degenerate sense. To see the obviousness of the former, we have only to recall the familiar fact that in most (almost all) normal cases, even if there is little or no continuity of memory and character etc, we do make identity-judgments about (non-contemporaneous) persons provided bodily continuity is assured and that we are generally right (and others agree with us) about it. To put it in another way, if the two criteria point in opposite directions, as in the case of amnesia or permanent loss of memory, it does not even occur to us that the person after the psychological change might be different from the person we see before the change. Moreover, the rationale of the "who has suffered the change?" question supplies the logical foundation to our saying that "they" are one and the same person. It is, perhaps, significant that normally there is thought to be no "problem" in the described cases. Why then should there be any "problem" if the criteria seem to conflict in the reverse order, i.e. if, in spite of bodily discontinuity, continuity of memory and character is displayed with considerable accuracy? Can't these cases, few and rare as they are, be explained as cases of what has been described as retrocognition, clairvoyance or some such para-normal experience? (If

it is not too strong\* a way of putting it, we may also describe such cases as rare pieces of coincidence, i.e. of x's seeming to remember doing a and it being the case that a was done in the past.) With the availability of these sorts of explanations which supply a clear analogue to amnesia and such like in the "unproblematic" cases, the conclusion that persons do survive their bodily death (and sooner or later become resurrected or reincarnated), and hence that memory is the sole criterion of personal identity, does not seem to be clearly mandatory. Moreover, we have already seen that, in the absence of bodily continuity, such conclusion would be unwarranted and also that, for the same reason, even if we talked of "same person" in such unusual circumstances, this would only have a secondary, degenerated sense. However our suggested treatment of the conflicts of criteria would seem too simplistic unless we said something fully and specifically as to how the so-called "puzzle cases"<sup>1</sup> are to be dealt with and what, if any, relevance these cases can have to the problem

---

\*One reason why it would be too strong to call it a coincidence might be that such cases may keep repeating. But if something (or some type of things) repeats itself some of the times, that itself does not make it any better than a coincidence; e.g. if someone on several occasions (or several people on several occasions) said that a certain thing is going to happen and it did happen, we will not grant that he (or they) knew it - not, especially, if there is a distinction between knowledge and "lucky guess" (which we take to be no different from coincidence). In our language, we make room for what are called 'exceptions', and the rare cases of these types can have this one explanation. Of course if things of this sort keep repeating more often than not, our inclination may be to regard them as cases of knowledge simply on the basis of the run of their success and accuracy; but this will mean a considerable breakdown of our conceptual system - with "knowledge" assigned a different logical role from its usual one. As Wittgenstein expressed this, if things did not happen as they normally do and if exceptions became the rules and rules exceptions, then we would not be playing the language game that we do play. It would at best, be a different game or no game at all.

<sup>1</sup>I would like to call all cases of criteria-conflict "puzzle cases" so that the rather unproblematic cases like amnesia and paramnesia would come under that description, and so would the supposed phenomena of resurrection and reincarnation. However, for convenience of historical reference, I shall reserve the label to mean those "problematic" cases like "bodily transfer", brain-transplant and split-brain transplants etc.

of personal identity. For one thing, it is these cases which have been devised and stressed to demonstrate the conflict of criteria by showing how the joint application of the criteria leads us to puzzles. And for another thing, it might be said that in the cases which we've just considered above (as cases of ~~criteria~~ conflict) the criteria do not really conflict but only that one of the criteria (memory, in the case of amnesia etc and bodily continuity in the case of survival) is in-applicable. (Although, if what we've said in the last two sections is right, the inapplicability of the bodily criterion will discredit the case of reincarnation and such like cases).

Now then, without needing to repeat the list of the "puzzle cases" it can be said without much disagreement that these cases have been usually invented to suggest one of the two following things: (1) that in such cases there is no right answer to the question of who is who, and (2) that these cases show that memory is the sole criterion of personal identity. (Keeping the meaning of "sole criterion" as ambiguous and unclear as we have shown it to be, see sec ii of the last chapter.)

Now, to take suggestion (2) first. Although this has been the purported suggestion, explicit or implicit, in all traditional and most contemporary interpretations of the "puzzle cases", we have already seen that this suggestion is unwarranted. For what these cases are capable of showing, it has been argued, is either that the continuity of memory (and character etc) is due to the continuity of a bodily part, e.g. the brain (see Ch 3 sec ii ) or that the memory-like claims in such cases being real remembering is highly dubious and as such its role as criteria is questionable (see last section). This applies to the standard type of "puzzle cases" like the Lockean-Quintonian cases of "bodily transfer" and to the Shoemaker-type cases of "change-of-body" caused by brain transfer. (And as our account of survival and disembodied existence would have made clear, this would also apply to the

alleged cases of resurrection and reincarnation as well as to the possibility of "life-prolonging" process by means of recurring brain transplants). But this will also apply to the other type of puzzle cases e.g. the Dr Jekyll/Mr Hyde type of cases where onebody is supposedly inhabited by different persons. For although the memories in such cases will not be questioned as in the other cases, these cases will still not show clearly that there are two persons rather than two different personalities being displayed in the same person.<sup>1</sup> (And the fact of there being two different persons being thus in question, it will not have been shown that memory is the sole criterion of personal identity in any sense.) Indeed, as will be argued shortly, the different (or conflicting) streams of memories might be due to the one and the same person having had different (or conflicting) series of experiences in his past. Wittgenstein, in exploring the possibility of a similar "puzzle case", asks us to imagine

"... a man whose memories in the even days of his life comprise the events of all these days, skipping entirely what happened on the odd days. On the other hand, he remembers on an odd day what happened on the previous odd days, but his memory then skips the even days without a feeling of discontinuity,"

and contends that it would be neither right nor wrong to say that there are two different persons rather than two different personalities in the same person (id. p 62), and, by implication, that "we could say whichever we like" (cf p 62, few lines earlier). Besides falling straightforwardly into a rather odd view of persons and personal identity (namely, that the question is a matter of decision\*), Wittgenstein's view here seems to be based on an inadequate analysis of the imagined situation. For on one

---

\*We shall argue later that with adequate qualifications, the oddity of such a view could be avoided.

interpretation at least it will appear that, not only are we not "bound to say" that there are two persons in this case, but also that it would be wrong to say that, and this for the following reason: Unless we assume that personal identity is to be defined in terms of memory, what the present case shows is that the memories of the person comprised of two different series , one concerning the events and actions of the odd days of his life and the other concerning <sup>those of</sup> the even days of his life; and the apparent alienation between the two series of memories (one series containing no element at all of the other and vice versa) might possibly be explained to be due to the person's having had (being subjected to) alienating experiences on the odd and even days of his life. The last point may be made out by imagining the following thought experiment. Right from the dawn of consciousness in him, a child is brought up alternately in two entirely different surroundings with entirely different people and things around him. Let us suppose that on every odd day of his life he is in surrounding A and that on every even day he is in surrounding B; let us also suppose that at the end of each day he lapses into a short coma during which he is shifted quickly from the one surrounding to the other. As a result of this, let us further suppose, the person displays two entirely different sets of memories, character and personality etc, like Wittgenstein described in his imaginary case. In this case, I believe, it will be certainly wrong to say that there are two persons living in one body, for saying this would imply the absurdity of saying that we could, as it were, "manufacture" two (or any number of) persons out of one original person by subjecting him to our described procedure. And since our case is, in principle, no different from Wittgenstein's, it follows that in the latter too it will be wrong to say that there are two different persons inhabiting one and the same body. Thus, far from showing the primacy of

the memory criterion, this "puzzle case" has all the potencies - which Wittgenstein did not quite realise - of pointing against it. As we have already seen, the temptation (to say that the puzzle cases show the primacy of the memory criterion) is more obviously mistaken in the case of what we called the "scientific thesis" of change-of-body; in particular it has been seen that memory continuity which, in these cases, would seem to ascertain personal identity would be due to bodily continuity - either in the form of brain continuity or, if we like, in the form of a brain being merely a catalytic factor in the exercise of memory and consciousness in the human body. The same will also hold, mutatis mutandis, in the much more complicated case of split brain transplant which - if at all it shows identity - will not show the primacy of the memory criterion. (The last case has been discussed fully in Ch 3 seciv and also in Appendix 2).

We have thus seen that, of the two supposed implications of the "puzzle cases" (2) above is definitely false. Does it follow from this that (1) is true? Does it mean, i.e. that we have no right answer in such cases? This has been the general attitude of several contemporary writers<sup>2</sup> towards the puzzle cases, although some also have wanted to resist saying this.<sup>3</sup> In view especially of this difference of opinion, this suggestion has to be qualified and understood in its possible perspectives before being accepted or rejected. The suggestion that there is, in the puzzle cases, no right answer might be understood in a rather extreme sense to mean that it would be meaningless to say one way or the other (since, presumably it could be equally right or wrong to say the one thing rather than the other). This, we have seen, might be the intention of Parfit when he said, in the split-brain transplant case, that we cannot say that one of the 'products' was the same person as the original person (the brain donor).<sup>4</sup> We have argued that unless meaning and verification were taken to be the same, and even meaning and criteria

were the same, this will not be the right thing to say. For even if nothing would "show", or no criteria could be there for saying, that only one of them is the same person as the original person, the latter might be true nonetheless. We have also seen, in our analysis of reincarnation and resurrection, that the alleged survivor might intelligibly be said to be the same person as the "deceased" he claimed to be, though this would be said only in a secondary sense. In the similar vein, it will not be straightforwardly unintelligible (much less self-contradictory) to say that the prince/cobbler case or the Bultitude case are cases of bodily exchange. However, what we have said, and the least we want to maintain, is that we have no justification in making identity-judgments in any of these cases. And this may be a second sense in which the suggestion in question can be interpreted, namely that in the puzzle cases there is no justified answer to the question of who is who. If our analysis of the puzzle cases here and elsewhere before is correct, then this seems to be the right solution (though our analysis has also revealed the logical balance swinging in favour of bodily continuity and, so, of non-identity and no change-of-body). With no ground to say in either way, you could say "whichever we like", and this is true not only in the Dr Jekyll/Mr Hyde-type cases, but also in the alleged cases of bodily exchange or bodily transfer. However, saying this has certain other implications which need careful discerning. Saying this means that, in these cases, the question of who is who is "a matter of decision" which must be distinguished from two awkward implications: (i) that it must be an arbitrary decision and (ii) that it must be a verbal decision so as to have the effect of changing our concepts of a person and personal identity. (People who wanted to resist saying that there is no right answer in the puzzle cases were apparently (partly at least) trying to avoid these last implications.)<sup>5</sup>

On the lines in which we have outlined our account of personal identity, it is possible to say that there is no right (justified) answer in the "puzzle cases" without committing ourselves to these last implications. For we have argued that bodily continuity and memory are the criteria of personal identity and given our definition of "criterion" (See Ch 3 sec.1 ) the puzzle cases pose no real difficulties for our account of personal identity. They only show - quite in keeping with our view - that in certain cases, abnormal as they are, the criteria fail to give us any clear result. The puzzle cases are puzzle cases, and what makes the element of "puzzle" plausible is the (natural) fact that our criteria fail to guide us in these cases. It is but natural, therefore, that judgments of identity/non-identity must be withheld, or kept under suspense, in these cases as any such judgment will have no criterial basis and so no justification. However, if in these cases we have to say one way or the other, the issue would depend on a decision; but the decision will not have to be an arbitrary one, since reasons can be given for settling the identity-question one way rather than the other: reasons like what the law-courts would decree, or what the family, friends and people associated with the victims would say, and so on - reasons, that is, which are more or less of an objective nature. It is important here to note that what these considerations point to is what it would be best to say in the case, which is quite another thing from what actually is the case - and even from what justifiably can be said to be the case. Therefore, this type of solution to our puzzle situations will only be a matter of expediency rather than any real solution worth the name. This underlines the significance of our saying that there is no right answer in these cases. However, even if we are willing to allow the issue to be thus a matter of decision, the decision need not be a "verbal decision" as to how the words "same person" are to be used. In the latter case, we will have to change the

concept of personal identity by revising or re-defining the concept: and those who wanted to do this were anxious to do this in order to accommodate the puzzle cases in their general account (definition) of personal identity. They have wanted to say that a correct account of personal identity must not only apply to the ordinary cases but also to the puzzle cases, so that the assertion of personal identity in those cases may have the same logical status as it has in the ordinary cases. And this is because they wanted to define personal identity; and a definition that suited the puzzle cases was adopted - even if it embarrassed the ordinary notion of personal identity, even if it threatened to change the ordinary concepts of a person and personal identity. Thus, Locke defined personal identity in terms of memory (his word was "consciousness") because "should the soul of a prince carrying with it the consciousness of the prince's past life enter and inform the body of a cobbler ..." etc, the latter will be the same person as the prince and vice versa; (and we have seen what oddity it creates for our normal concept of personal identity as it is related to the issue of moral responsibility, see Ch 2 sec i ). And Parfit has offered an entirely new concept of personal identity which even required us to give up the language of identity because this concept, and not the normal one, applies to the puzzle cases (fission and fusion and split-brain transplants which most interested him). But what they were defining personal identity in terms of was only a criterion and not, as we've seen, even a necessary condition. Since we have also seen that the concept of personal identity cannot be defined, the futility of such attempts as the above need not be over-emphasised. On the contrary, our analysis of the "puzzle cases" above should have shown that these cases create no conceptual difficulties as these theorists seem to have thought; and therefore that there is no need to change or even extend our normal

concept of personal identity in accordance with what we may say in these cases. For as will be pointed out shortly (and see also the last section), contrary to the apparent assumption of these theorists, assertion of personal identity - if made in these cases - will not have the same logical status as it has in the normal cases; and if so, these assertions fall far short of supplying a paradigm for the general application of "same person". But before going into this, it will be worthwhile here to point out (in a rather general way of theorising) that the difficulty the problem cases create can at best be said to be a crieterial difficulty which cannot be a conceptual difficulty, unless criteria and meaning were the same (which we have argued not to be). If the normal criteria do not apply (or be of help) in some odd cases this cannot change the general meaning of "same person" - nor even can it discredit the normal criteria or make them cease to be the criteria. The existence of "puzzle cases" could not threaten our normal concept of personal identity and the criteria thereof any more than the existence of clones and amoebas threaten Leibniz's Law.<sup>6</sup> It can at best pretend to call for the adoption of new or modified criteria only to suit these cases.<sup>7</sup> (However, our reservations, which follow shortly, to this suggestion must be noted carefully.) But we will not be obliged to change (or even extend) the meaning of "same person" simply on this ground. For firstly, as is consequent upon what we have said above, change of criteria does not necessarily mean a change in the meaning of the relevant concept. And secondly, as we have argued in the last chapter. (see especially p 242, 244-5 ) these so-called "new" criteria will fall far short of their crieterial role. They will not certainly be criteria if the puzzle case is one that belongs to what we have described as the "naive thesis" of alleged change-of-bodies. For there, in the absence in principle of any independent bodily check, the "memory claim", which we are told should

be the criterion of identity, could never be known to be real remembering and as such would only be an apology for a criterion: it would, at best, be a memory-like-claim, but never a memory-claim. Thus, in effect, our making of personal identity judgments in such cases would be based on no criterion (but only on what looks like one). This, incidentally, would explain our contention that the meaning of "same person" in these cases, as in the alleged case of disembodied survival will have to be different from our normal understanding of this concept: we might be having a person in the normal sense of the term; but the idea/claim that he is the same person as some earlier person (he might claim to be), being based on a degenerate criterion as above, will have to be degenerate itself. Nor can it be said to follow that the new or modified criteria are the standard criteria any more than that the new concept - if one is innovated - would be the standard concept of person and same person.\*

It can be said, further, not only that we are not obliged to change our normal concept of "same person" in view of the odd nature of the puzzle cases, but that, because of the very nature of the identity-assertions in these cases, it would be wrong to do so. For as has been argued in the last two sections of the previous chapter, <sup>and</sup> reiterated in the above paragraph, with no possibility of checking the genuineness of memory-claims, the assertion of identity on the basis of mere memory-like-claims alone gives "identity" only a secondary, degenerate sense. And since this is precisely what we do in the puzzle cases (except perhaps, in the cases of brain-transplant and fission and fusion, which have their own peculiar qualifications and difficulties<sup>1</sup>), our talk in

---

\*This argument belongs to the family of arguments by which we have shown that the concept of dis-embodied persons can't be the standard concept of persons so as to make the Cartesian hypothesis of non-embodied existence intelligible.

<sup>1</sup>See Ch 3 sec 4v.

these cases, of "same person" would have a secondary, degenerate sense. We can't therefore, replace our normal concept by this concept. For, as was shown, the secondary use of a concept is intelligible only because of the existence of the primary use (which is our normal use of "same person"); but with this use giving place to the secondary use suitable to the puzzle cases, no intelligible sense can be given to the latter.

Further, that the talk of "same person" in the puzzle cases would not be of the same logical status as our talk of the same in the normal cases will be shown by the fact that this will not have any justification whatsoever whereas our talk of "same person" in the normal cases is sufficiently justified by the satisfaction of the criteria of bodily continuity and/or memory. If there is any difference between speaking with justification and speaking without any, then the assertion of identity in the puzzle cases would be of a very odd sort indeed. Moreover, we are not forced to choose between the two answers - in terms of identity or non-identity - in the puzzle cases. We could, in fact, say neither and, taking all the facts into consideration, there is nothing absurd in starting with a "new" person - a Brownson for example.\* This person remembers what Brown did because he has Brown's brain; he looks like and has most, if not all, physical features of Robinson because he has the latter's (brainless) body. But that is all; henceforth he will have different experiences which neither Brown nor Robinson had anticipated, will remember doing all sorts of actions which neither of them had done, and so on. Given this reasonable choice, our talk of identity in these cases would seem to be all the more peculiar and so of a very

---

\*With appropriate wordings and qualifications, a similar "third" person description will not be difficult in the prince/cobbler or Bultitude case.

different nature from our talk of this in the normal cases.

I conclude, therefore, that the puzzle cases pose no conceptual difficulty to the issue of personal identity, and, therefore, that even if we may have to choose between one or the other of the answers as to whether or not the earlier and the latter persons involved in these cases are the same, it does not necessitate nor justify any change or extension in the concept of "same person". They would pose a conceptual difficulty if the problem were a problem of meaning or definition which, we maintained, it is not.

Finally, now what relevance the puzzle cases have to the problem of personal identity? After outlining the way, we have, as to how these cases are to be dealt with, it would look very natural to claim that these cases have no relevance to the problem. In view of the reasons given above, I am inclined to believe that the spirit of this claim is essentially true - though its letter needs some clarifications so as to avoid misunderstanding. By our claim that the puzzle cases are not relevant to the issue of personal identity is to be understood no more and no less than the following two claims: (1) that they do not justify any change or extension in the meaning of "person" and "same person" and (2) that they do not justify change or extension even in our present criteria of personal identity. (What they may, of course, lead us to do is to adopt some new or modified concepts like "retro-cognition" (cf. Penelhum) or some new or modified criteria\* (Shoemaker) in order for us to be able to describe - if we must - these situations in identity terms. This, in effect, will be a sort of "conceptual innovation",<sup>8</sup> but an innovation only to suit these cases, and what is more to the point, those new concepts or criteria will be as vague and

---

\*But see pp 265-66 above and note 7 now.

unclear<sup>9</sup> as the situations they would be meant to describe. Moreover, since we have shown that we are not obliged to describe these cases in terms of identity/non-identity (with any earlier person) the described conceptual innovation will not be necessary. And even if we are forced to take resort to this, by certain pragmatic considerations or some such thing, that will leave everything as it is; nothing will happen to our current practice - neither the concept of personal identity nor the criteria thereof will be touched. Nevertheless, the importance of the puzzle cases can hardly be undermined. It will not be denied that it is the existence/possibility of puzzle cases that makes personal identity a philosophical problem. And this it does by seeming to disprove what ordinarily is believed to be true, namely that bodily identity plays at least the central role in the assertion of personal identity. But seeming to disprove something is one thing and disproving that thing is quite another. And although, as we've shown, the puzzle cases do not disprove what their authors wished them to, they are still of philosophical interest and will remain so, in so far as they make us study carefully the concept we have of person and of same person. In the face of these cases (or in anticipation of their possibility) we are made to analyse and articulate the knowledge we have of persons and personal identity, and to uncover and tighten the principles that are implicit in our judgments of personal identity (it needs a philosopher to explicate the law of non-contradiction - though the common knowledge of it is reflected in our ordinary talks and acts. But then, the philosopher "leaves everything as it is", he does not change matters but "only in the end describes them" (Wittgenstein)). This process of articulation and analysis could have shown our current practice to be mistaken; but in the present case, as we have seen, contrary to the expressed intention of most authors of the puzzle cases, the analysis

has pointed rather in the opposite direction. It has supplied our knowledge, which perhaps unreflectingly used to pass as such, with a theoretical basis. It has, for example, clarified (so far, at least, as we reckoned it) that the familiar central role of the bodily criterion should be understood, not as a defining condition, nor even as a necessary condition, but as a criterion of personal identity - albeit an independent and the primary criterion in the sense we've given to it. If not for this, it is in initiating the discussion of these various alternative perspectives of our common knowledge that the main philosophical importance of the "puzzle cases" must lie. In view of this, the problem of personal identity would hardly be a philosophical problem without the occurrence/speculation of these cases. But, as we have said while this is one thing, the truth of what the puzzle cases purport to show is quite another.

Section (ii): Must There Be An Answer? - The importance of personal identity

We have seen that it is mainly because of the possibility of "problem cases" that a fresh review of the nature of personal identity as well as its importance is called for. And I have suggested that all this review amounts to is an analysis of the concepts of person and personal identity - and not a revision thereof, as most writers seem to have thought. For, it has been argued, the possibility of these cases - and even the adoption of new concepts and/or criteria to suit these cases - is perfectly compatible with our normal concepts of person and personal identity remaining as it is. To recall the conclusion of the last section, "problem cases" do not create any conceptual difficulty to the issue of personal identity (so as to call for a conceptual revision). If this is so, there are two remaining questions which we will have to consider by way of final analysis: first, why do we demand an answer, - an all-or-nothing answer, - to the identity questions about persons; and secondly, why - as is suggested by many analyses of the problem cases - is it that such answers are not always possible.\* (Roughly speaking, the first question concerns the nature of personal identity and the second is more concerned with the issue of criteria thereof.)

When we say that a later person is the same person as an earlier person, we mean that the two persons are identical simpliciter. And we say this and mean this even if the grounds on which we say this are (and are known to be) only imperfect indicators<sup>1</sup> of identity. Accordingly,

---

\*It will be recalled that when we say that no answer is possible in such cases, all we mean is that no justified answer is possible. This must not be confused with the other claim, which we do not make, that there could be no true answer in these cases. We maintain that there can be a true answer but only in the sense that any answer may happen to be true - though, we insist, we will have no right to say which one is this.

when we are faced with two non-contemporaneous persons and the question of their identity is at stake, we naturally expect a yes-or-no answer to this question. Usually, there are thought to be two reasons: one is a belief about the nature of personhood and personal identity, and the other is what is thought to be the importance of personal identity. The first belief is that personhood and personal identity is a deeper fact - that it is something more than the bodily and psychological facts which it inevitably involves, so that despite the variant nature of these latter facts we do not feel obliged to think or talk of personal identity holding to varying degrees - even though our making of personal-identity judgments is grounded on these facts. Historically, the most explicit theoretical expression of this belief (the "Simple View", as Parfit calls it) is to be found in the writings of Reid and Butler who described personal identity as "perfect identity" despite the (bodily and psychological) evidences thereof being admittedly "imperfect". What this "something more", according to these philosophers, is unanalysable and can only be revealed to us through some sort of intellectual intuition. Perhaps, like its nearest kin, the Cartesian "privileged access" theory of personhood, this formulation of the belief in the special nature of personal identity is in a sense indefensible.<sup>2</sup>

But what I want to suggest is that the weakness of this formulation of the belief does not rule out its plausibility in any other formulation, and does not necessarily yield place to anything like a Parfitian theory of degrees of identity. (It must be recalled here that although I do not agree with this Cartesian theory of personhood nor explicitly with the intuitionist theory of personal identity, yet I do specify the sense in which the nature of personhood and personal identity has a speciality of its own as distinct from that of other material objects;

see Ch 1 sec ii.) Presently, I will try to bring out the plausibility of the "Simple View" by expressing it in a different but acceptable way, and show the importance of demanding all-or-nothing answers to personal identity questions. But let us mention, in passing, the second reason for demanding such answers. This reason consists of the belief that personal identity is presupposed by certain other important questions like memory, survival and moral responsibility,<sup>3</sup> and that since these latter questions usually demand all-or-nothing answers so does the question of personal identity. Let us now leave this belief to be discussed and defended later on after we have given a plausible analysis of the first belief.

Perhaps it is slightly misleading to state the "Simple View" by saying that personhood and personal identity is a deeper fact, for as indicated above, with this metaphysical undertone (especially with its family resemblance to Cartesian mentalism) the theory may more easily appear to be indefensible. But it is not necessary to throw away the baby along with the bath water. I suggest that the truth of the matter lies in the way we speak of persons and personal identity; and so a proper linguistic analysis of these concepts and their application-rules will reveal the truth more accurately than any supposition of what we may believe about their nature. As we use the word "person" (i.e. as we apply the word to whatever thing or being we do) the word has what may be called a unitary function. That means that, by virtue of whatever qualities or properties the word is applied, a person - once he is so called - is as much a (fully-fledged) person as any other person, i.e. as any other being who is so called.<sup>4</sup> Each person stands for a single unit no matter whatever is the extent or degree of the qualities he/she may be judged to have - and this applies even to those qualities that may be central or essential to one's being a person. A more rational man is

said to be as much a person as any other man who may be much less rational. Centuries after Aristotle defined man as a rational animal,\* no one has ever thought of such things as a 'half man' or a 'double person' (as distinct from double personality), not to mention "mini-person" or "maxi-person" and the like, because of, and despite, the familiar fact that people are more or less rational. Consequently, whatever it is - call it 0 - that makes someone a person can never affect his personhood through its own imperfections. In other words, if x's are persons by virtue of their having (or being) 0, no x is more a person than another x, even if he has (or is) more 0 than that other x.  $x_1$  and  $x_2$  will be two persons (unless of course they are identical) even if  $x_1$  has (is) double (if that can be measured) the 0 than  $x_2$ ; and  $x_1$  as well as  $x_2$  will have one single vote if they are asked to exercise their franchise. Incidentally, those who reason and act according to this principle need not be and seldom are, guided by any consideration like personhood being a deeper fact, but they are simply guided by the way we speak/and "pick out" <sup>of</sup> persons and the way we use the word "person". Thus, in our linguistic practice, we have learnt to use "person" as an all-or-nothing word; any x either is or is not a person, and if he is he is as much a (fully-fledged) person as any other x who is a person. And for this purpose, no two x's need to be similar (much less exactly similar) in respect of bodily and/or psychological features, even if these features are essential to an x's being a person. It is, I think, in this sense that personhood is, in its logic, all-or-nothing. By similar consideration of our usage of "same person" we reach the same conclusion about personal identity:

---

\*It will not do violence to the spirit of the present argument to ignore the rather dubious Lockean distinction between person and man, and to take 'man' and 'person' as synonymous or interchangeable. Aristotle, apparently, did not use "man" as distinct from "person" in his definition, and so I will use them interchangeably here. (For a reasoned rejection of the Lockean distinction see Wiggins "Stream of Consciousness" in Philosophy, April 1976 p 142, note 23)

two non-contemporaneous persons either are or are not the same person as each other, and if they are, they are so despite there being nothing unchanged between them and irrespective of their bodily and psychological continuities (which may be all that matters for personal identity - but more of this, later) being minimal or maximal. Anyone who has not been introduced to the idea of a "something deeper" about the nature of person-hood can be taught and can quite easily master the technique of using "person" and "personal identity" as efficiently as anybody else. For it will be enough for him to know and observe that this is the way the words are used. The idea of "something deeper" is thus not only an unnecessary appendage to the understanding of these concepts, it can also be a misleading factor in this understanding as the history of the subject may testify.¶ And I think, it is this misleading feature of the idea that Parfit, consciously or unconsciously, utilised in his campaign against the "Simple View".) The supposition of "something deeper" can only multiply entities beyond necessity, and in as much as the "Simple View" is wedded to this supposition it must have its failings. But we can see from the foregoing, the "Simple View" can be freed of this supposition and yet have its say. If this is so, there is no need to hasten to make a revisionary proposal as to how we ought to use our concepts of 'person' and 'personal identity' in order to replace the "Simple View".<sup>5</sup> In all three of his papers on the subject<sup>6</sup> Parfit has attacked the "Simple View" basically and expressly, though not only, on the grounds of the "something deeper" supposition. And in view at least of the superfluity of this alleged supposition (as we have shown above), his criticism must be ill - founded.

We have just indicated above how the logic of the words "person" and "personal identity" requires all-or-nothing treatment of personhood and personal identity. This means that, if we know the use of these words well enough, we will not think that there can be more and less to the nature of persons and personal identity. Of course, a person can be more or less something (viz intelligent, strong, capable of remembering and so on) but as a person he can't more or less. And

similarly, there may be more bodily and/or psychological continuity between  $P_1$  and  $P_2$  than between  $*P_1$  and  $*P_2$ ; but for that matter  $P_2$  is not more identical with  $P_1$  than  $*P_2$  is with  $*P_1$ . It is the oddity of the expressions "more identical" and "less identical" - something that concerns the logic of these words - which rules out, as likewise odd, the nature of personal identity being a matter of degree. Since the last analysis of the nature of personal identity (and person-hood) conflicts with the logic of these words, our analysis above suggests that this must be mistaken - that something must have gone wrong with the Parfitian analysis. And indeed something has gone wrong - this in the following way: It is one thing to say (a) that what matters in  $\phi$  is  $\psi$  and quite another thing to say (b) that  $\phi$  is just a matter of  $\psi$ . A very familiar illustration may be the distinction between saying, as we might, (a<sup>1</sup>) that what matters in (knowing that some-one is in ) pain is that people generally should display what is called pain-behaviour and saying (b<sup>1</sup>) that pain is just a matter of pain-behaviour. It needs no great philosophical ingenuity to see that (b<sup>1</sup>) does not follow ( i.e. pain  $\neq$  pain-behaviour), even if (a<sup>1</sup>) is true. But the "even if" is of particular significance when the reasoning from (a) to (b) - fallacious as it is - is applied in the way Parfit would want to apply to the case of personal identity. For I shall categorically question the truth of step (a) in his reasoning. But now we see that, granting the truth of (a) in his reasoning (namely, that what matters in personal identity is/are bodily and/or psychological continuities), the (b) in his reasoning (namely, that personal identity is only a matter of degree) does not follow. It has been argued (see Ch 3 sec iv) that Parfit runs into this fallacious conclusion through the doubly mistaken procedure of supposing that the problem of personal identity is a problem of definition and of trying to define it in terms of what only are the evidences thereof. But furthermore, granting that (a) what matters in personal identity is the described continuities, how can be articulate(a)- what, i.e., do we mean by (a)? One way in which it can be interpreted is by saying, as

we have said, that these continuities are what make us apply the words "same person"(this will concern the logic of personal identity, since what concerns the rules of usage of a word/words must concern its/their logic - but yet we do not say, nor does Parfit, that personal identity is, in its logic, a matter of degree()). Another way in which (a) can be interpreted is by saying that this is what we know when we know pairs of persons to be the same as each other (i.e. we know personal identity by knowing that these continuities hold between these persons). This is what naturally will be thought to concern the nature of personal identity since this relates to the way we come to know what personal identity is. Parfit's conclusion about the nature of personal identity is certainly based on this way of articulating (a).<sup>7</sup> But if this is his ground, it is not clear how the conclusion that he wants to draw follows, how, that is, the nature of personal identity is a matter of degree. For one thing, although it is true that we know cases of personal identity by knowing these continuities, it does not follow that personal identity is what these continuities are - even if it were true that we cannot know anything else apart from these continuities in knowing personal identity. Consider the following. All we know\* when we acquire the knowledge of a table is simply that side ( or those sides) of the table which faces us, and it is impossible that we can know the other side (s) of the table or indeed anything else about it at the very moment when we know it inevitably from one side of it. Yet from this it does not follow that the (nature of the ) table is only its facing side. If the phenomenalist claims to acquaint us with the nature of physical objects by means of his sense-data analysis, the way of phenomenism must be mistaken. There is certainly more to the nature of the table than just the facing side of it or what we know or observe about it. And saying this does not commit one to anything metaphysical or mysterious about

---

\*Our use here of "know" in the sense of "see" or "observe" should be perfectly innocuous for the purpose of the present analogy; since this precisely is the sense in which we know the described continuities when we are said to know personal identity.

the nature of the table. Similarly, the "Simple View" would be certainly right in holding that there is more to the nature of personal identity than simply the bodily/psychological continuities which it involves or which we know when we know personal identity - and in holding this, it will not be committed to any metaphysical or mysterious implications about the nature of personal identity (as Parfit's purported emphasis on "something deeper" seems to suggest). It follows therefore that Parfit needs further arguments to justify the alleged separation between the logic and the nature of personal identity,<sup>8</sup> since the point of this separation is not sufficiently clear from what he provides. Moreover, the truth that the nature of personal identity is not to be analysed in terms of the described continuities - even if it were true that the nature of  $\phi$  consists in how we know  $\phi$  - is evident from the fact that knowing these continuities is not the only way of knowing (the nature of) personal identity. For self-identity (personal identity of one's own) is not known in this way (See Ch 1 sec ii and Ch 2 sec ii), and it has been argued before that self-identity is what, in a more important sense, acquaints us with the real nature of personal identity and that the knowledge of this identity can be, and often is, non-criterial. So that the continuities, in an important sense, become fairly dispensable to the nature of personal identity.

Now, we are led straight on to consider, for whatever it is worth, the (a) in Parfit's reasoning. How far is it true to say that what matters in personal identity is the described continuities? Our last consideration, namely that concerning the nature of the knowledge of self-identity makes this claim rather dubious; and at any rate, our above analysis of the nature of personal identity as related to how we know it, has shown that these continuities are not all that matters - even though they do matter a great deal - for understanding the nature of

personal identity. Therefore, a one-to-one correspondence in the relation between these continuities and the nature of personal identity (viz, since the former is a matter of degree, so must be the latter) is out of the question. Furthermore, Parfit's own statement of, and belief in, the nature of the claim (a) makes it all the more dubious and farther from truth. For as we have already noted, he believes, and claims more emphatically, that what matters in personal identity is only "psychological continuity" and bodily continuity is totally and unjustly ignored.<sup>9</sup> But it is not true that what matters in personal identity is only psychological continuity; indeed, if what has been said in the previous chapter is correct, bodily continuity plays a more fundamental role in knowing and making judgments about personal identity. What then were the grounds for thinking that psychological continuity is all that matters or even mostly that matters in personal identity? One stock argument which Parfit makes use of is that psychological continuity is morally important - important, that is, for ascribing moral responsibility<sup>10</sup> (and therefore that personal identity which is presupposed by the question of moral responsibility must be analysed in terms of this continuity). We have little more to add here than what we have said before, about the weakness and fateful consequence of this theory which, like its traditional ancestor, the Lockean theory of "consciousness of guilt", is "morally repugnant",<sup>11</sup> to say the very least. What I wish to add here is that even if we ascribe moral responsibility solely on the basis of psychological continuity, that will not make personal identity a matter only of psychological continuity. And even if it is undoubtedly true that beliefs about the nature of personal identity are relevant to the ascription of moral responsibility,<sup>12</sup> the converse does not hold. For even if we accept the "complex view" and excuse a person on grounds of (apparent) psychological discontinuity

("marked change" in character, style of life and loss of memory etc) that will not show that he is a different person from the accused, for one thing at least, he is excused because it is he who has changed in the described respects\*. Yet curiously, Parfit eminently in his first two papers seems to rely on this reverse reasoning (from how we ascribe moral responsibility to what is the nature of personal identity) in order to establish his "Complex View". As a second strand of argument for his favoured claim, Parfit cites the examples of (1) what we do say and (2) what we might say about our earlier (as well as latter) selves. On the first score, he quotes from Proust<sup>13</sup> and Solzhensitsyn<sup>14</sup> to show that we already have a use for the distinction between "successive selves", which is made on the basis of lessening of psychological connection. (But surely, as it is evident from Parfit's illustration of this, this use is a figurative use or poetic use; and no argument has been supplied to show that this is, or should be, the right or proper use.) But as Smart has rightly shown,<sup>15</sup> what this talk, if taken seriously, will show is not that a person is a series of selves (so that they would be different selves connected only by psychological continuity), but that the person is (in the sense of identity, and expressed tenselessly) each of these selves: these "different" selves are only the phase-restrictions on one and the same person, and as such, are identical with each other, and cannot be different from each other. Thus even if we do say (and say seriously) the above about the successive selves, it will be simply mistaken to say, on that ground, that psychological continuity is all that matters in personal identity. On the second

---

\*And note, in this connection, Mackie: "Why should we not also say, 'He did it, but he does not remember doing it' rather than 'He does not remember, so it was not he that did it'". (Problems from Locke, p 196)

score, (2) above, i.e. on what we might say about our so-called earlier and later selves, Parfit says,

"When, for instance, we have undergone any marked change in character, conviction, or style of life, we might say, 'It was not I who did that but an earlier self'"<sup>16</sup>

and apparently takes this as a ground for showing the importance of psychological continuity. Now, besides being subject to the same objection as his claim about what we do say, there is a sense in which this second claim can be said to be false. For it can be argued that if someone said this, it does not show that he has undergone any "marked change" of character etc, but rather quite to the contrary. Especially, if it is responsibility for an evil deed that is at stake, and a person says what Parfit suggests he might say, then that will more likely indicate that his vicious character has not changed and that being still prompted by this nature he is only trying to avoid responsibility. (Is it not what in fact is happening in most criminal proceedings?) Thus this suggestion that we might "non-identify" our present self from an earlier self can't show these "selves" to be different from each other, nor that psychological continuity is all that matters in personal identity.

Let me now summarise my conclusions of the present section, and then we proceed. These conclusions are (1) that it is our usage of the words "person" and "personal identity" - the logic of these concepts - that makes us believe in the all-or-nothing nature of personhood and personal identity, and demand all-or-nothing answers to the identity questions about persons. And (2) since what matters - if that matters - in personal identity, namely, psychological continuity, has little to do with the nature of personal identity, it is mistaken and needless to try to prize apart the nature of personal identity from its logic and to describe the former, in terms of that continuity, as a matter

only of degree. However, (3) the alleged importance of psychological continuity is not only ill-founded\* but also entirely mistaken,<sup>/</sup> and therefore the nature of this continuity being a matter of degree has got nothing to impair the importance of and the all-or-nothing nature of personal identity; consequently we are not obliged nor justified to change to the "Complex View" which, being based on this mistaken assumption, is mistaken itself.

To these conclusions, and particularly to the last, we will also add the conclusion of an earlier chapter (Ch 3 seciv) about the importance of personal identity. (And here we come specifically to the second reason as mentioned above for the belief in the all-or-nothing nature of personal identity.) For after proposing his "Complex View" Parfit had questioned the importance of personal identity on another ground. He had argued that personal identity is considered important because it is presupposed by certain other important questions like those of memory, survival and moral responsibility, but, he argued, these latter questions can be freed of this presupposition and personal identity can thus be made to lose its importance. We have shown the futility of such a proposal by arguing that with personal identity losing importance, these questions of memory, survival and responsibility - will be of no importance themselves. Since the latter will be an unacceptable consequence, the importance of personal identity must remain beyond question. Indeed, for a proper understanding and use of the concepts of memory, survival and responsibility as well as for a just and coherent theory of moral responsibility, the importance of personal identity

---

\*Evident from my contrary suggestion as to why we might "non-identify" from an earlier self. (See last page).

<sup>/</sup>Evident from Smart's phase-sortal analysis of the successive selves. (See last page).

is an essential pre-requisite. Hence the importance of all-or-nothing answers to personal identity questions. This conclusion may be reinforced by means of yet another procedure. This procedure is to restore the primary importance of identity by bringing the importance back from the notion of "survival". In ordinary usage the notion of survival is almost inseparable from the notion of identity and the latter is often articulated in terms of survival and persistence conditions.<sup>17</sup> But Parfit's proposal is to detach the two notions of personal identity and survival and to place the primary importance on the latter. But as Smart has argued,<sup>18</sup> it is not clear and no argument has been supplied as to why this is so? It is not clear from what point of view "survival" is so important, and "what, precisely, is supposed to survive?" Let us, on our part, try to make it clear by considering different possible senses in which "survival" can be effected. Smart suggests two: the first is what he calls "personal survival" and the second, "survival of something personal" (op cit p 25). We can express this by saying that survival can be of (i) me and of (ii) my powers (like my memory, my character, my projects and so on). Although (i) is the ordinary notion of 'survival', Parfit opted for (ii) by throwing light on the fact that even if I do not survive, the survival of my powers and projects can be still important to me. Indeed, to take Smart's example, a composer might regard the survival of his powers to be more important than his own survival: "his brain might be duplicated and someone else could write the music that the original composer would have written were he able to survive" (op cit p 25). Thus, Smart concedes, connectedness<sup>ness</sup> guarantees the survival of my powers. However, it does not seem to be a very satisfactory notion. For it is not made sufficiently clear whether a surviving power will be mine rather than someone else's power that is exactly like mine. It is

conceivable that after I cease to exist, someone else might come up with exactly similar abilities as mine to compose exactly the same kind of music as I used to compose and would have composed. In view of this possibility, I shall add a third sense in which the notion of survival may be articulated, namely (iii) the survival of powers like mine.\* And since in the mere event of the emergence of certain powers, there is nothing to tell the distinction between (ii) and (iii), we will have no reasonable ground to say that (ii) has happened rather than (iii); and the prospect of survival in terms of (ii), though important and attractive to some, will not be to much purpose. For the person himself can't have reasonable ground to rejoice over such a prospect, and this for the following reason. As Williams has shown, the result of brain-duplication, information-transfer etc - even where they are carried out - is highly dubious as to whether it is going to be me who survives.<sup>19</sup> By a slight stretch of Williams' arguments, the same can be said about the question of whether it is going to be my powers. For even if the success of a brain-cuplication would guarantee the survival of my powers, it cannot be said or expected with reasonable certainty that it will be a success, that what happens after the operation is the result of the success of the operation rather than that, by a perfect coincidence, some entirely new powers (though they are exactly like mine) which belonged to someone else have come to survive in this new brain/body complex. If the fact or rather the possibility is as I have described, then the only reasonable way to articulate and anticipate survival is in the form of (i) i.e. the survival of me. But this is inseparable from

---

\*There will be nothing, in the nature of the "surviving" powers, to guarantee the form of survival as (ii). And my basic purpose in introducing the third form of survival is to diminish the import of (ii) by arguing that there would be no reasonable ground, even in the event of brain-duplications and the like, to tell the difference between (ii) happening and (iii) happening.

the idea of my personal identity. Therefore the only reasonable account of "survival" is that which makes "survival" and "identity" interchangeable. And since Parfit's project purports to reject this (only) reasonable sense of "survival", the project must be rejected. Thus, if the above analysis is right, then what is of primary importance is personal identity, and hence the importance of all-or-nothing answers.

Finally, we can now return briefly to our original second question of why it is that such answers do not always seem to be possible. We have argued that personal identity is, both in its logic and in its nature, all-or-nothing. Consequently, whenever an answer in identity-terms is possible, it must be in the form of a Yes or a No. The trouble seems to arise, however, in the "problem cases". But what exactly happens in these cases? The most that can be said about these cases is that there seems to be no right answer since any answer either way will have no justification. We have suggested in the last section that, in view of the extreme abnormality of some such cases, we are not obliged to give an answer in terms of identity or non-identity, that we can talk of a "third" person beginning to exist anew. But even if we are obliged, because of certain practical considerations and the like, to say whether or not the person after the change ~~is~~ the same person as the person before the change, we will only be judging that they are completely identical or completely different. Whatever considerations incline us to say that Locke's cobbler-bodied person is the same person as the prince and vice versa, we will say that they are identical simpliciter, i.e. in the all-or-nothing sense. Any suggestion that, in these cases (or in some very complicated ones), identity should be ascribed in terms of degrees, has been shown to be unwarranted and needless. And the corresponding suggestion that our concept of personal identity should be redefined to admit of degrees is to be rejected.

"Problem cases" are problem cases and cannot creat conceptual difficulties to the issue of personal identity. Nevertheless, under some very compelling circumstances and/or because of equally compelling considerations we may be obliged or inclined to adopt a modified criterion or at least to modify the modus operandi of our present criteria (e.g. to make identity judgment solely or primarily on the basis of memory\*). But this will only show that our criteria are not suited to give the right result in all possible cases. And in view of our account of the nature and function of criteria, this difficulty is absolutely unworrying. For criteria are justificatory evidences, and no definitions, nor even necessarily necessary conditions; and if they (or their normal conjunction) fail to guide us in certain circumstances, it is only to be expected. To be sure, our understanding and use of concepts depends a great deal, often essentially, upon what we use as criteria for their application. And although criteria and meaning are not the same, yet any change in the criterion is likely to affect the nature of the concepts; for we may pick out the wrong object, or may not even know what to pick out, if the criteria for the application of the concept are changed or absent. It is in this sense that meaning and criteria are closely connected. Yet it will be unreasonable to expect that our criterion should be able to guide us in all possible circumstances. For, as Wittgenstein said, "We do not use language according to strict rules - it has not been taught to us by means of strict rules, either."<sup>20</sup> If it were, and if criteria were able to guide us in all possible cases they would not be

---

\*Though we have argued already that such a procedure will inevitably lead to a secondary and degenerate use of the concepts of person and personal identity. And this at any rate is an argument that, these cases - intelligible and plausible in whatever sense they are - cannot provide a standard for defining or describing the concept of personal identity.

criteria (but defining conditions and would necessarily entail, which a criterion need not do, see Ch 3 sec i. above). Similarly, understanding the use of our concepts is to be able to use them in all normal circumstances; but this is not to expect that we will be able to apply the concept (and answer questions concerning its application) confidently in all possible cases. On the contrary. To understand a concept is not only to be able to answer questions concerning its application in all normal circumstances and to answer them correctly in most of these cases, but also to be doubtful in the cases which are doubtful.<sup>21</sup> This is particularly true of our problem cases where, knowing all the facts as we do or can, we still may not know what to say. But once again, this is entirely unworrying and can neither affect our normal concepts of person and personal identity nor the criteria thereof.

Appendix 1

SPECIAL NATURE OF PERSONAL IDENTITY  
AND THE IDEA OF "PRIVILEGED ACCESS"

§1 A very central part of my thesis, as advanced in the foregoing, has been that persons have a special way of knowing their identity through time and that this accounts for the special nature of the problem of personal identity as distinct from that of identity of other things. And this speciality, it has been argued, lies in the fact that every person can have non-criterial knowledge of his own identity whereas his knowledge of the identity of other things has to be grounded on criteria. In expounding this contention, I do indeed speak of a "privileged access" that persons have to the knowledge of their own identities. But this should not be misunderstood. In particular, my position does not commit me to entertain any form of spiritualism or Cartesian mentalism. For this belief in the "privileged access" to the knowledge of oneself led the Cartesians to a peculiar belief about the nature of persons - i.e. that they are essentially minds or non-physical entities, whereas my analysis of the "privileged access" is confined only to the nature of the knowledge of personal identity, to the fact that personal identity is (or can be) known in the way it is. And this fact has no (any special) consequence on the nature of persons. My position can be vindicated on the following ground: side by side claiming that the knowledge of personal identity can be non-criterial I have also maintained that there are empirically applicable criteria for the across-time identity of persons. Indeed this latter forms the life-blood of my treatment of the problem of personal identity (namely, that it is a problem of criteria). And as, I hope, is more than implicit in my thesis, the link between criterion-independent self-knowledge or self-ascription<sup>1</sup> and the publicly usable criteria of subject-identity is not "in practice" severed, though in a logical analysis of the nature of

self-knowledge I have claimed it to be. For I have, in the thesis, made adequate provision for the fact that, even in one's own case, it is imaginable that one may be unsure, or even mistaken, of one's identity and so be prepared to accept correction from others who inevitably have to use criteria for telling him who he is. It is with this important pre-requisite (i.e. the linkage with the empirical criteria of identity) that self-ascriptive utterances or judgments can refer to the person (the subject of these ascriptions). But if we forget this and emphasise only on the non-criterial nature of self-knowledge, we fall into the Cartesian illusion of supposing that the subject of our self-ascriptive utterances is a purely inner, immaterial substance (cf Strawson, Bounds of Sense, pp 165-66). As Strawson rightly pointed out, this delusive notion of the person (or the "I") is the result of severing its connection with the empirical concept of the subject of experience (cf Id p 166). As we have taken due note of the criteria of personal identity and are basically concerned with them, we are doing nothing of this sort. The "privileged access" we speak of is a logical privilege concerned with the analysis of the nature of self-knowledge, rather than an ontological privilege regarding the nature of persons. It is in view of this that when, for instance, we said, in vindicating the Reid-Butler view, that (the knowledge of) personal identity is not vitiated by the "imperfections" of criteria (see pp 45 and 80 above) nothing like a sense of piety or profundity need be attached to this. It only points to the brute fact that this knowledge is acquired in the way it is.

§2 Given this account of the "privileged access", it puts the person himself in a better position to say or know who he is, but it does not make him infallible.\* (Nor is it necessarily the case that the only things

---

\*Compare Ryle: 'The superiority of the speaker's knowledge of what he is doing over that of the listener does not indicate that he has Privileged Access to facts of a type inevitably inaccessible to the listener, but only that he is in a very good position to know what the listener is often in a very poor position to know' (my emphasis). The Concept of Mind Penguin 1973 p 171

that can be known non-criterially are the so-called private experiences. For I can know, without using any criteria, that I am writing an appendix to my thesis, and this latter fact is never a private event inaccessible to anybody else.<sup>2)</sup> As I have said, in certain circumstances, others may have a better say on this. Yet if, in certain peculiar situations, others cannot clearly agree as to who a person is (as in Shoemaker's Brownson case) the person himself may be given the last word on this (see the end of sec ii of Ch 1 above) in as much as what others say stands the risk of being falsified by what he says. And what he says will depend surely on what he remembers, since, by supposition, there is no bodily continuity here; and this may have the pretensions of making memory the sole criterion of personal identity. But as our treatment of the problem cases (see esp Ch 5 sec i) would have made clear, even if, in the case in question, we grant the last word to the person himself the above result will not follow. For (i) the supposed case is not a normal case, so what we may be forced to say in this case will not affect our normal concept of a person and of personal identity, and what we may be using as our ground for making identity judgment about this person will not affect our normal criteria; and secondly (ii) only if we were interested in defining personal identity then saying what we say in this case would oblige us to define personal identity in terms of memory - but this clearly is not our concern and so we are in no fear of being wedded to a memory theory in the described way.

Appendix 2

SPLIT-BRAIN TRANSPLANTS AND  
MEANINGFULNESS OF PERSONAL IDENTITY JUDGMENTS

So far as the issue of personal identity is concerned, the possibility of bisected brain transplants will have three alternative explanations. For if such transplants succeed and are successfully accompanied by an apparent continuity of memory and character etc in the two "products" of the transplant, then we may say (1) that both the products are identical with the original person (the brain donor), or (2) that only one of them is or (3) that none of them is. As in my treatment of the general possibility of reduplication (e.g. the Charles/Robert-Guy Fawkes type cases) so in this case, I am inclined to accept (3) and reject (1) and (2). The explanation (1) is unacceptable since each resulting person, supposedly being the same person as the original one, has to be the same person as the other resulting person; and this, for well-known logical reasons, is an incoherent consequence (see pp 156-57 for my argument for this) - unless, like Parfit, we are prepared to say that identity does not matter (see Ch 3 sec v for the unacceptability of the latter suggestion). The second alternative will have two variations: (2-a) that one of the resulting persons is identical with the original person if the other does not survive or is not allowed to survive the operation (i.e. if, for example, one of the half-brains does not take<sup>1</sup>), and that (2-b) even if both the resulting persons continue to exist and are found to be similar in all observable respects (viz, in respect of memory, character, personality etc), yet only one of them is or may be identical with the original by virtue of his having some additional features which we may not or cannot observe.<sup>2</sup> Now (2-a) can be easily dismissed on the ground that this would boil down to the absurd implication that who one is depends on what happens to someone else!

(That is, if I am the product of a bisected brain transplant, then I am the same person as the brain donor if my "competitor" does not exist or is prevented from coming into being, but I am not if he does.) Now with (1) and (2-a) being thus dismissed, (2-b) seems to be the only plausible explanation. It is plausible because although we cannot observe any additional feature which one of the products may have, it does not follow that there is no such additional feature, and so it does not follow that one of the resulting persons is not or cannot be identical with the original person. To say that the last follows is to insist, as Parfit apparently did (discussed in Ch 3 section iv), that one of the "products" must differ in respect of some observable features in order for him to be identical with the original person. And it may be said, as indeed it has been,<sup>3</sup> that this is to rely on the verification theory of meaning, that unless a statement is confirmable with reference to some observation (observation-statement(s), to be precise), it must be meaningless. If this is so then (2-b) can be definitively dismissed provided the verification theory is established on strong grounds. I have no intention of defending this theory as I do not think that any very good reasons can be adduced in its favour. I am, however, inclined to believe that (2-b) cannot be a satisfactory solution to our problem. For one thing, to say that only one of the "products" is identical with the original person would at least be vacuous - if not meaningless. It will be vacuous in the sense that an identity-judgment cannot be justifiably made to this effect. This is particularly significant for our purpose, since we have maintained that the problem of personal identity is one of criterion (which, in our account, is a justificatory evidence), and hence no solution to this problem will be satisfactory if it does not, and cannot, specify any criterion for making personal identity judgments. In view of this, it may be perfectly meaningful and

, coherent to suppose that only one of the resulting persons is identical with the original person (i.e. by virtue of some supposed additional feature which he may have), but it will not be a helpful supposition to make and will not throw any light on what grounds can we have to say that a later person is the same person as an earlier - much less on what it is on which personal identity consists in, if that is what concerns some people. (And since these are the only two ways in which the problem of personal identity can be raised, a solution on the model of (2-b) will be no solution.) Further, as is evident from the foregoing, the supposition expressed by (2-b) may be true, but the trouble with it is that it cannot be known to be true. On the contrary, there would be overwhelming empirical evidence that it is not true. For, ex hypothesi, each half-brain carries with it the same memory and character etc, in fact, all that is supposedly necessary and sufficient for its recipient to be the same person as the brain donor; and this is what supposedly makes the resulting persons exactly similar<sup>4</sup> (and also it is this supposition that gives the present problem the sort of perplexing form it has - i.e. whether each is, or none is, or only one is the same person as the original person). It will be, therefore, most unlikely that any one of the resulting persons will have an additional feature which will make him, and not the other, the same as the original.

My position, then, is fairly clear. I do not want to say, nor am I committed to say that since there is nothing observably special about the one person rather than the other, (2-b) must be meaningless; and thus I am not committed to any form of verificationism (though, as I said, I should be happy with the success of the latter). But I do maintain that (2-b) will be unsatisfactory: saying that one of the resulting persons may have some additional features (which we know not what) and as such may be the same person as the original person will not say anything significant

or useful - neither about the nature of personal identity nor about the grounds on which identity-judgments may be made about persons.

Incidentally, this type of explanation will not be of any help to defend Parfit's position from the similar charge of verificationism. For he seems to be implicitly committed to the latter, and that for the following reasons: I have pointed out, while dealing with his theory of personal identity and elsewhere, that Parfit takes the problem as one of the meaning of personal identity, and is concerned with what personal identity is or in what does it consist. And his answer to this question is that it consists in nothing but the observable evidences that we have that a later person is the same person as an earlier person. For he insists that the fact of personal identity consists in the holding of certain more specific facts which, presumably, are the facts like bodily continuity and the continuity or similarity of memory and character etc. (He refers to the latter as "psychological continuity" and unreasonably insists on this alone - to the total exclusion of bodily continuity - as constituting personal identity). Apart from these "specific facts" he does not believe in any further fact that may constitute personal identity. But if the nature of personal identity is thus to consist in these specific facts then, in the absence of these facts a later person cannot be the same person as an earlier person. Further these specific facts are what we can observe to hold between persons at different times; and if the meaning of personal identity is to consist in these and nothing else then his commitment to verificationism becomes fairly apparent. The inescapable outcome of such a theory would be that in the absence of these observable facts there can be no personal identity and any claim that there is or may be is meaningless. The additional feature which (2-b) alludes to and which, we have seen, may be there (at least it is not unreasonable to suppose so) in one of the "products" of a

bisected brain-transplant, will not be an observable feature (for if it were, we'll know it); it will, therefore, not be these "specific facts", not even anything like them. It would follow, therefore, that on Parfit's account, the supposition (2-b) is meaningless. As we have said, this is verificationism which Parfit must defend before his theory can attain the status of acceptability.

NOTES TO THE TEXT

Chapter - 1: Section(1).

1. T.Penelhum "Hume on Personal identity" in Philosophical Review (hereafter PR) 1955, reprinted in David Hume ed V.C.Chapell, pp 213 (Macmillan, London 1968)
2. See G.Vesey, Personal identity (Macmillan, London 1974)
3. This other area involves the rather trivial issue of something's being identical with itself. The question of such identity are discussed non-trivially when the thing in question comes under different descriptions even at the same instant of time, e.g. when we say that the flash of lighting which I saw is the same flash of lightning which you saw, or the more philosophically interesting issue that the 'Morning Star' is the same star as the 'Evening Star'. It will be agreed that since there is no specific question of reidentification of things at different times our present problem will not be concerned with this. For the same reason also, I will not be discussing "identity across kinds" like the question of whether mental states are(the same as) the brain states.
4. E.g.A.J.Ayer, The Concept of A Person and Other Essays (Macmillan, London 1963) Ch 1 and also his The Problem of Knowledge (Pellican paper backs,1956) Ch v secs i and ii. See also Vesey Personal Identity, esp his distinction between the Unity- and the Identity-Questions.
5. Cf J.F.M. Hunter, "Personal Identity", in his Essays After Wittgenstein (Allen and Unwin, London 1973) p.35.
6. Thomas Reid, Essays in the Intellectual powers of Man(hereafter Essays) ed A.D. Woodzley, (Macmillan, London 1941), P 200.
7. Hospers distinguished various senses in which such question may be asked in his Introduction to Philosophical Analysis, revised

edition (Routledge and Kegan Paul, Lond 1967 paperback 1973)  
pp 12-13.

8. By a definition being non-trivial I mean that it should not be trivial for those who already know what the definiendum means. In this sense what I have given as examples of trivial definitions are trivial to us, to those, that is, who are fairly well-versed in the language containing the words "same" and "different". These definitions, though, may be quite illuminating (hence not trivial) to a foreigner or to children, to those, i.e. who are beginners in the learning of the language. The last, however, is a different matter.
9. See Locke, An Essay Concerning Human Understanding (hereafter Essay), BK II Ch 27 sec.9.
10. Essays p 202
11. The example and the argument has been taken from Penelhum. See his "Personal Identity" in the Encyclopaedia of Philosophy ed. P Edwards Vol VI, and also his "Hume on Personal Identity".
12. See esp P.T. Geach, Reference and Generality (Cornell University Press Ithaca N.Y. 1962) Sections 31-34 and ch six; and also D Wiggins, Identity and Spatio- Temporal Continuity (ISTC hereafter) Basil Blackwell, Oxford 1967).
13. Cf Wittgenstein, e.g. in his Blue Book (Blackwell paperback 1972) esp. p 61.
14. Swinburne op cit pp 231-2.
15. This has been notably argued by Bernard Williams in his "Personal Identity and Individuation" PAS 1956, reprinted in his Problems of the Self (Cambridge University Press 1973).
16. See Swinburne op cit esp p 234 ff.
17. See Swinburne, op cit p 241.

18. Bishop Butler, "Of Personal Identity" in his The Analogy of Religion (hereafter Analogy) (George Bell & Sons London 1902) p 238; Reid, Essays p 202.
19. J.L. Mackie, Problems from Locke (Clarendon Press, Oxford 1976) esp pp 192-3.
20. Sydney Shoemaker, Self-Knowledge and Self-Identity (Cornell University Press, Ithaca, N.Y. 1963) p 5.
21. Ibid p 3.

Chapter - 1: Sec (ii)

- 1 David Hume, A Treatise of Human Nature (hereafter, Treatise) Everyman's Library, London 1911, Vol I p 245.
2. "Hume on Personal Identity", in David Hume p 231
- 3 Essays, p 205.
- 4 Wittgenstein once remarked that if a lion could talk we will not understand him. (Cf. Philosophical Investigations, Basil Black well, London, Part I sec 223, also sees 19 and 23).
5. This, in fact, is a fallacy which is similar to what Perry had called the fallacy of "initial predication" and attributed to the idealist philosophers. (See Perry R.B. and others, The New Realism, Macmillan, N.Y. 1925, pp 15-16).
6. Strawson said that certain predicates like "is in the drawing room", "was hit by a stone" etc, might mean one thing when applied to material objects and another when applied to persons. P.F. Strawson, Individuals (University Paper back, Nethuen, London 1964) p 105.
7. Ibid p 89. The things which are ascribed both to ourselves and material things are the physical properties which imply that we

both have bodies; the things which are ascribed to ourselves and could not be ascribed to material things are what may be called the psychological features ( a further distinction between persons and some animals in this respect follows shortly in the text) which imply that only persons have minds. And nothing at all seems to follow from saying this to the effect that persons may or can have no bodies much less that persons are minds.

8. Essays p 201.
9. Essay Bk II Ch 27 sec 9.
10. Cf Philosophical Investigations.
11. See William James, Principles of Psychology, (Macmillan, London 1962) Vol I Ch.X .
12. Harry Frankfurt brings out this distinctive feature of persons, more convincingly, in terms of what he calls the higher order intuitions" which only persons can, and no animals can, be said to have. See his article "Freedom of will and The Concept of A person", in JP 1971.
13. Spatio-Temporal continuity, however, is of much wider extension, and spatio-temporal continuity of what should be taken as the criterion of identity - where it is a criterion - is a further, often debatable, issue. But what will not be debated is that the answer to this question depends largely on the nature of the thing (the concept which the thing comes under) that is being judged. Depending on what the thing essentially is (what the concept necessarily involves in order to have a use or application) the spatio-temporal continuity of the thing is ~~not~~ determined and identity ascribed. Some types of thing i.e. ships, would require continuity under a common form as the criterion for their identity, x while some other type of objects, e.g. watches, would require the

spatio-temporal continuity of parts as their identity criterion (for a fuller discussion of this see Brian J. Smart "How to Reidentify the ship of Theseus" Analysis, 1971-72 and "The Ship of Theseus, The Parthenon and Disassembled Objects" Analysis 1973-74). However, this consideration of details need not detain us here. For our purpose here is simply to point out that since spatio-temporal continuity (of some sort or other) is normally used as the criterion of identity for both material things and of persons, this might, in all probabilities, have led some to say that our knowledge of identity in both cases was based on the same ground.

14. Williams, for example, who champions the case of spatio-temporal continuity (of the human body) to be always a necessary condition of personal identity, admits that it is not a sufficient condition (See "Personal Identity and Individuation" Problems of the Self pl)
15. See Self-Knowledge and Self-Identity, Ch 6 secs 5-7.
16. Ibid, Ch 4 and elsewhere.
17. Ibid, pp 34-5
18. E.g, Reid, Essays P 204 and Butler, Analogy, cf pp 29-30.
19. Self-Knowledge and Self-Identity, p 38.
20. Essays pp 201-2. It is very surprising that Shoemaker should have overlooked this and given an erroneous interpretation of Reid. For even he himself has said (see Self-Knowledge and Self-Identity, p 36 para 1) that these philosophers "have held that the concept of identity is indefinable" ( my emphasis).
21. Analogy, pp 329-30
22. Ibid, p 330.
23. R Chisholm, "The Loose and Popular Sense and the Strict and Philosophical sense of Identity", in Perception And Personal Identity

proceedings of the 1967 Oberlin Colloquium in Philosophy, ed by  
N.S. Care and R.H. Grimm (Case Western Reserve, Cleveland 1969)  
P.82.

24. Ibid p 83

25. Brian J Smart has argued that in the described case it is the  
replaced ship - not the reassembled one - which should be the same  
ship as the original ship, since although 'identity of parts' could  
be a criterion of ship identity, yet the parts that go to form the  
reassembled ship ceased to be the parts of the original ship by virtue  
of having been not only removed from the latter but also replaced by  
the aluminium parts (see his article, "How to Reidentify the Ship of  
Theseus", Analysis April 72 p 148, and also his "Personal Identity in  
an Organised Parcel", Philosophical Studies , 1973 p 421). Wiggins also  
appears to have argued in favour of the same conclusion on the ground  
that there is, at every stage of the replacement, a ship ("continuity of  
form" criterion) that continues throughout the described process of  
replacement (see his ISTC, p 37). These arguments (which, implicitly  
and explicitly, decree in favour of the 'continuity of form' criterion as  
opposed to the "identity of parts" criterion)- sound though they are -  
do not have anything to show the crucial importance of these criteria  
to the meaning of "ship". For even if we would be persuaded, on these  
grounds, to say that the aluminium ship is the same ship as the original  
ship, we will not thereby mean anything different by "ship" now from  
what we understood by it earlier. By contrast, as we will see shortly,  
our acceptance of the psychological criteria to the exclusion of the  
physical criteria (and conversely) will, in each case commit us to a  
different concept of person from that embraced by a rival contender.

26. Ibid .

27. Shoemaker attacks on the rather trivial ground that "since what counts as the same  $\phi$  is always determined, at least in part, by the meaning of the word  $\phi$ , there is a sense in which all criteria of identity are conventional" (Ibid pp 112-13) and that personal identity is no exception.
28. Ibid, p 115.
29. Ibid, p 115.
30. ISTC, p 57.
31. Cf, Shoemaker Self-Knowledge and Self-identity, pp 31-33.

Chapter 2 : Section (i)

1. See A.M. Quanton "On Punishment", Analysis Vol 14 (1953-4)
2. Especially sec 16, "Consciousness makes the same person" (my emphasis) and sec 19, "This may show wherein personal identity consists" (my emphases). Also see secs 9, 10, 23 and 29.
3. Ibid, II 27 Section 9 (emphasis mine)
4. A.G.N. Flew, "Locke and the Problem of Personal Identity", Philosophy, 1951.
5. Analogy, p 329.
6. See Williams "Personal Identity and Individuation" in Problems of the Self.
7. See Reid, Essays, pp 213-14.
8. It might be thought that the awkwardness of this consequence can be avoided by the Gricean-Quantonian theory of continuity according to which though all the experiences in a personal history cannot be remembered continuously, it would be enough if memory serves to form a continuous series of interconnected experiences. As it was expressed, it would be enough if each phase of experience is

directly connected to its proximate phase ( in virtue of containing memory of the latter) and, through it, indirectly connected to the more remote phases. Mackie gives a similar explanation in terms of what he calls "memory bridges" which, connecting the unconscious phases of our lives to our conscious ones, might constitute "the sort of continuity that is appropriate to persons" (Problems from Locke, pp 180-1). It will, however, be borne in mind that this sort of theory is a revision of Locke's theory of personal identity and is indeed a better and more plausible account of it. But the really interesting thing is that Locke never intended to say what this theory says about personal identity, and in fact he was committed explicitly to a different view (See Mackie Ibid, p 181). He would even go on to invoke the distinction between "same person" and "same man" to explain the school boy/ officer / general paradox by saying that although the general is the same man as the school boy, he is not the same person (cf Essay, II Ch 27 sec 20). He would embrace this counter- intuitive, if not absurd, position just because he wanted to define personal identity in terms of (direct) continuity of memory or consciousness, and no such definition could cope with a situation like the described one. Now, the revised version of Locke's theory might be able to cope with the situation satisfactorily, but this it can do only because it is not an analysis of personal identity in terms of continuity of consciousness (As Mackie admits "openly" in his book, see pp 192 and 196), but presumably as a factual analysis - which is nearer to my criterio-logical analysis.

9. See Treatise, Vol 1 p 245.

10. Ibid, p 251,

11. Ibid, p 252.

12. Ibid, p 262.

13. Chisholm, "On the observability of the Self", Philosophy and Phenomenological Research, '69-70.
14. D.F. Pears, "Hume on Personal Identity" in David Hume - A Symposium ed. himself, (Macmillan, London 1963).
15. op cit, p 53

Chapter 2: Section (ii).

1. Bishop Butler, Analogy.  
Thomas Reid, Essays.
2. See Ch 1 sec 1 above.
3. See Analogy p 332.  
  
The Lockeans, Butler argued there, cannot consistently be claiming that persons are really the same, since consciousness which they say constitutes personal identity, is not the same individual act at all times. And he goes on to add, "The bare unfolding this notion, and laying it thus naked and open, seems to be the best confutation to it".
4. Essay, Bk II, Ch 27 secs 11 and 12.
5. See Shoemaker, Self-Knowledge and Self-Identity, . . , p.46 and also Reid, Essays pp 212-13.
6. Essays , . p 200.
7. Analogy, p 332.
8. "Memory and Personal Identity", AJP 1964.
9. op cit.
10. Analogy P.330.
11. Ibid, p 330.
12. Ibid, p 330 (the following paragraph)
13. Ibid, p 332.

14. Of course, speaking without justification is not to speak without meaning. But why do we say that a soul continues in a person ? Surely, to invent something to meet the requirement of "something unchanged" to account for our identity through time, and this we have seen to be an unnecessary requirement. It follows, therefore that the making of personal identity judgments in terms of (or by supposing) the continuity of a soul is not only unjustified but also needless.
15. E.g. by Shoemaker, Self-Knowledge and Self-Identity, Ch I and also pp 254-58.
16. Reid, Essays p 214
17. Butler, Analogy p 329
18. Evident from: "But to say that my remembrance that I did such a thing or my consciousness makes me the person who did it is .... an absurdity too gross to be entertained by any man who attends to the meaning of it" (my emphasis) Reid Essays p 214.
19. Ibid, p 204.

Chapter 2: Section (iii).

1. Critique of Pure Reason, trans, by N.K.Smith (Macmillan, London 1923). The numbered references hereafter in this section preceded by A and B are to the page numbers of the first and the second editions of the Critique respectively.
2. The conclusion of the third paralogism was that the Soul is "a person" (A 361). About this conclusion, at least, Kant has no disagreement with his relationist adversaries, though where he differs is about the explanation of the identity of persons.
3. I shall leave out the claim of the 4th paralogism and Kant's

comment on it as I do not think it comes under the scope of our discussion here.

4. A 363.
5. The Bounds of Sense (Methuen, London 1966), p 162.
6. B 413. Compare Hume: "When I enter most intimately into what I call myself, I always stumble on some particular perception of other ... I never catch myself at any time without a perception, and can never observe anything but perceptions" (Treatise I, Pt IV Ch 6 ).
7. B 413.
8. Kant quoted from his Anthropology by W.H.Walsh in his Reason and Experience (The Clarendon Press Oxford 1947), p 196.
9. Ibid, In the same passage from Kant.
10. See Chisholm, "On the Observability of the Self".
11. A 362 .
12. J Bennet, Kant's Dialectic (The University Press Cambridge 1974), p 100.
13. Ibid, p 95.
14. Quotated by Walsh, Reason and Experience, p 18.
15. For a fuller discussion on this point, see Walsh, Ibid, Ch 9.
16. Kant even admits this at one stage, though very obscurely. (see e.g. B415).
17. Strawson, The Bounds of Sense, p 164.
18. The bracketed numbers here and hereafter in this section refer to the page number of James Principles of Psychology (Macmillan, London 1902), Vol I Ch x, Sec on "Pure Ego".
19. A.J.Ayer, Origins of Pragmatism (Macmillan London, 1968), p 265.
20. That it is not self-contradictory is pleaded, among others, by Strawson. See his Individuals Ch 3.

21. James had acknowledged in his Pragmatism that the idea of his pragmatism and even the word "pragmatism" was suggested to him by the Critiques of Kant.
22. Treatise, Appendix.
23. Origins of Pragmatism, p 266

Chapter 3 see (1).

1. "How can persons be ascribed M-predicates ?", Mind, January, 1977.  
An earlier ~~xx~~ version of this paper was presented at the staff and Post-graduate seminar of the Keele University Philosophy Department.
2. J.H.Hick, "Theology and verification", partially reprinted in Flew(ed), Body, Mind and Death (Collier-Macmillan, London, 1964).
3. J.L.Pollock, "Criteria and Our Knowledge of the Material World" PR Vol 76 (1967), p 32.
4. W.G. Lycan, "Non-inductive Evidence: Recent Works on Wittgenstein's Criterion", American Philosophical Quarterly (APQ, hereafter) (April 1971).
5. "Wittgenstein's Use of the Term 'Criterion'" in Wittgenstein, ed G. Pitcher (Papermac, Macmillan, London 1968), pp 243-4.
6. See Norman Malcolm, Knowledge and Certainty (Prentice Hall, Englewood Cliffs, N.J. 1963), "Do the propositions that describe the criterion of his being in pain logically imply the proposition, 'He is in pain'? Wittgenstein's answer is clearly in the negative" (p 113). And Malcolm goes on to say. "A criterion is satisfied only in certain circumstances" (P 113), and "the expressions of pain are a criterion of pain in certain surroundings, not in others", (p 114).

The last remark of Malcolm may be slightly misleading and so may seem Wittgenstein's views since the latter did say some-

thing to that effect, Cf PI, pt I see 584); for it has the apparent effect of claiming that a criterion is a criterion only in certain situations and not, presumably, in all. But in my view this rather hazardous way of expressing the matter has been taken resort to only to bring home the point that the c-relation is not one of entailment. A more correct reading of Wittgenstein may be approached by taking the "certain surroundings" and "other surroundings" as analogous to "normal circumstances" and "abnormal circumstances" which we have been discussing. The fact is that, it is the satisfaction of the criterion - not the criterion itself - which is said to depend on these circumstances. As our earlier analysis would have shown (so I hope), the truth rather is that it is the circumstances - whether or not they are to be called "normal" - that depend on or are relative to the criterion, and not the other way round.

7. The Blue and Brown Books (Basil Blackwell 1958), p 25. (my emphasis )
8. See esp, Ibid, p 25, where he x says that no strict line can be drawn between criteria and symptoms, and also PI, Pt I sec 354, where t he speaks of "fluctuations between criteria and symptoms".
9. Self-Knowledge and Self-Identity, p 3. This notion is elaborated and consistently applied in the book.
10. Lycan, op cit, p 110
11. Cf, op cit, p 110 two paragraphs earlier.
12. Routledge and Kegan Paul, London 1959.
13. Self Knowledge and Self Identity, p 4

Chapter 3 sec (ii)

1. More has been said about this distinctive feature of persons and the consequent speciality of the problem of personal identity in Ch 1 sec ii.
2. Originally published in the PAS 1956 and reprinted in his Problems of the self. Hereafter the page references to this article as well as

- to the others included in this book will be to the book.
3. Such theories have been suggested by Shoemaker and Parfit, among others. We will have occasion to consider them later on.
  4. Ibid, p 10
  5. Ibid, p 11. The plausibility and force of this argument seems to lie on the fact that the satisfaction of the memory- criterion does not show that bodily continuity is not there but merely that we do not have any direct evidence that it is. On the contrary, if the two persons in question were known to have different bodies, then despite the impressiveness of their psychological similarities, judgment of identity will be fairly suspect - if not withdrawn.
  6. Problems of the Self, p 10.
  7. In his "Are Persons Bodies ?", "The Self and the Future" as well as "Personal Identity and Individuation"; all are reprinted in the book referred to above.
  8. See "The Self and the Future", Ibid, p 63
  9. Cf, op cit, p 47 and also "Are Persons Bodies ?", Ibid, p 79.
  10. "The Self and The Future", Ibid, p 47.
  11. See Don Locke, Memory (Macmillan, London 1971), pp 135-37, where he has adduced such an argument for the reliability of memory.
  12. Essay Bk II, Ch 27 sec 16.
  13. A.M. Quinton "The Soul" JP, 1962
  14. See especially "Are Persons Bodies?" Problem of the Self, p 77
  15. Ibid, p 79.
  16. Ibid, p 78 also p 77.
  17. Ibid, p 77.
  18. Self Knowledge and Self-Identity, p 24.
  19. Ibid, p 24
  20. Ibid, pp 24-25.

21. J.M.Shorter "More About Bodily Continuity and Personal Identity",  
Analysis 1961-62,p 81
22. "Bodily Continuity and Personal Identity", Problems of the Self,p24.
23. Ibid,p 23.
24. "Personal Identity and Individuation" Ibid,pp 8-9
25. See Personal Identity, p 81.
26. Ibid,p 86
27. See "Bodily Continuity and Personal Identity" Problems of the Self,  
p 20.
28. Ibid , p 24.
29. "Personal Identity" PR 1971,p 13 (For a fuller account of  
Parfit" theory see sec iv below).
30. Personal Identity, p 90.
31. "On the Importance of Self Identity", JP 1971, p 689.
32. See Personal Identity,p 90

Chapter 3:sec (iii).

1. "Personal Identity and Individuation", Problems of the Self,p 1.
2. Compare Brian J.Smart, "Personal Identity in an Organized  
Parcel", Philosophical Studies ,1973,p 422" ... being an actual  
part (Which in the present context is the whole body-minus-brain,  
PKM) does not entail being an essential part."
3. "Are Persons Bodies?", Problems of the Self,p 77
4. See Wittgenstein, Blue Book,p 61; see also J.M. Shorter,  
" More about Bodily Continuity and Personal Identity",Analysis  
1961-62, pp61-62 and 81.
5. See, particularly, Wiggins ISTC ,p 37. See also Brian J.Smart,  
"How to Reidentify the Ship of Theseus", Analysis Vol.32 (1971-72).
6. See Brian J Smart "The Ship of Theseus,the Parthenon and  
Disassembled Objects", Analysis Vol 34 (73-74),p 25.

7. Self Knowledge and Self-Identity,p 24  
(emphasis added in the following quotation)
8. See Ibid,esp pp 174-5.
9. Ibid,p 194, also p 182 .
10. Ibid,pp 181-2
11. His review of Shoemaker's Self-Knowledge and Self-Identity in Philosophical Quarterly 1965,p 79
12. That Shoemaker was committed to this position as well is evident from his earlier paper "Personal Identity and Memory" (JP 1959), where he discussed this type of case as possible cases of "bodily exchange."
13. op cit,p 871.
14. Ibid,and also "Personal Identity and Memory", JP 1959.
15. Self-Knowledge and Self-Identity,Ch 4 secs 1 and 5 and elsewhere; also in "Personal Identity and Memory", op cit esp p 873.
16. Ch 1 sec (ii ) above.
17. Cf Shoemaker, "Personal Identity and Memory", op cit p 875.
18. Self-Knowledge and Self-Identity,p 243 Cf also "Personal Identity and Memory," op cit p 878.
19. See Self-Knowledge and Self-Identity,esp pp 246 and 245. Also "Personal Identity and Memory", op cit,p 879.
20. Thus he says,"By a criterion of identity for fs I mean something logically constutive of the identity of fs and potentially analytical of what it is to e an f", ISTC,p 43.
21. Ibid,p 55.
22. It has been pointed out in the last section that the brain being the bodily part that plays the role of the seat of memory and consciousness is only a contingent fact and that any other part or parts of the body could be imagined to play this role.Wiggins

gives vent to this idea especially in Ibid,p55 (few lines earlier to the quoted passage).

23. Cf. "Personal Identity and Individuation",Problems of the Self,p 12.

24. Cf. "Bodily Continuity and Personal Identity",Ibid,p 21.

25. Cf. B.J.Smart,"How to Reidentify the Ship of Theseus",Analysis,  
April 1972,p147.

26. "Wiggins on Identity",PR 1970,esp p 543.

27. Problems from Locke,see esp p 200.

Chapter 3: section (iv).

- 1 . This theory has been put forward in his "Personal Identity"(PR 1971),  
"On 'The Importance of Self-Identity'"(JP 1971) and his "Later Selves  
and Moral Principles" in Alan Montefiore ed. Philosophy and Personal  
Relations(Routledge and Kegan Paul,London 1973).The thesis was also gen-  
erally defended in a radio talk with Professor Vesey (the talk is  
printed in the latter's ed Philosophy in the Open (The Open University  
Press ,1974.)
- 2 . "On 'The Importance of Self-Identity'",(JP 1971).
- 3 . op cit,p 685.
- 4 . Terence Penelhum,"The Importance of Self-Identity",JP 1971 p 671.
- 5 . "On 'The Importance of Self-Identity'",op cit,pp 684-5.
- 6 . See his "Personal Identity",PR 1971.
- 7 . Penelhum,"The Importance of Self-Identity",JP 1971 p675.
- 8 . See his "On 'The Importance of Self-Identity'",JP 1971 p 688.
- 9 . See ISTC,esp p 7.
10. "Later Selves...",Philosophy and Personal Relations,p 138.
11. Cf. Ibid,p 139.
12. The suggestion that Parfit relies on the Verification Principle I got  
from Swinburne's paper "Personal Identity"(PAS 1974).I am not,however,

comfortable with the passage Swinburne quotes from Parfit to show this. In this passage Parfit considers Wiggin's example of a man's brain being split into two halves and each half being transplanted into a different (brazinless) body. He rejects the suggestion that only one of the resulting persons may be the same person as the brain doner on the ground that "each half of my brain is exactly similar, and,so,to start with,is each resulting person.So how can I survive as only one of the two people ? What can make me one of them rather than the other ?" (PR 1971,p 5). The assumption that one of them must differ from the other in some observable respect in order that he be the same person as the original person may be implicit here; but there seems to be no assumption - implicit or explicit - to the effect that the suggestion is meaningless.On the contrary,one sees Parfit saying (in the same page,few lines later) that the suggestion is "highly implausible" (which by no means is the same as "meaningless") and that it is one of the three"possible descriptions" of what can be said about Wiggin's operation.

Nevertheless it is quite evident that Parfit was working on some such assumption as that of the Verifi cation Principle,as I bring out in the text.

13. See Swinburne,op cit,p 231.
14. Cf. Malcolm,Dreaming (Routledge and Kegan Paul,London 1959),p 60, for similar arguments to show that criteria and meaning are not the same.
15. Swinburne,"Personal Identity",op cit,p 242.
16. See JP 1971,p 686. See also PR 1971,p 25: "Since this connectedness is a matter of degree,the drawing of the/<sup>distinctions</sup>can be left to the choice of the speaker/<sup>and</sup>be allowed to vary from context to context".
17. Cf. PR 1971,p 25.

18. That Parfit has not shown this and yet has to show this is alleged by Vesey in his Personal Identity, p 63
19. "Persons and Their Pasts" APQ 1970, p 284.
20. "Later Selves ... ", Philosophy and Personal Relations.
21. Essay Bk II Ch 27 sec 18.
22. "Later Selves ..."

Chapter 4: Section (i).

1. Locke Essay, Bk II Ch 27 sec 7.  
Shoemaker, Self- Knowledge and Self Identity, p 4.
2. See Ch 1 sec 1, for a fuller account of this.
3. More familiarly, in Williams' arguments in his "Personal Identity and Individuation", Problems of the Self.
4. See Williams, Ibid, p 10
5. Self- Knowledge and Self-Identity, p 1.
6. Read at the Keele University Philosophy Society on 12.2.76.
7. Cf Wittgenstein, PI, Pt I sec 282.
8. PI, pt I sec 282.

Chapter 4 section (ii).

1. See, for example, Shoemaker, "personal Identity and Memory", JP 1959, p 878; also Self-Knowledge and Self-Identity, Ch 6 section 8.
2. Cf Shoemaker, Ibid, Ch 6 sec 8.
3. Cf Ibid, esp pp 244-5.
4. Clarendon Press, Oxford 1974 (emphasis and brackets added).
5. That it is a case of approval is evident from his remark (few lines earlier, on p 26) that Williams' is "a more concise and coherent argument" and that "formally it is no better or worse than the Bultitude argument in that it consists in a straightforward

appear to our logical intuitions to what we should or should not want to say in certain circumstances."

6. A.M. Quinton, "The Soul", JP 1962, pp 404-5.
7. Cf Locke, "should the soul of a prince, carrying with it the consciousness of the prince's past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, every one sees he would be the same person as the prince ..."(Essay, II 27 15), and Quinton, "the soul ... is the essential constituent of personal identity" and "it is also what a person fundamentally is"("The Soul" JP 1962, p 403); my emphasis in both the quotations.
8. op cit, p 403.
9. Cf PI, sec 282.
10. Cf 'Personal Identity and Memory', JP 1959, p 878.

Chapter 4: Section (iii).

1. Cf Gilbert Ryle, Dilemmas (Tarnier Lectures, The University Press, Cambridge 1954), p 95.
2. See "Theology and Verification" in Flew (ed), Body, Mind and Death.
3. Penelhum, among others, considers such a possibility in his "Personal Identity, Memory and Survival," JP 1959. My reasons for disapproval of this theory follows in due course.
4. See e.g. Shoemaker, Self-Knowledge and Self-Identity, p 19.
5. St Thomas Aquinas, Summa Tehologica, Q 79 A2.
6. "Self, Mind and Body" in his Freedom and Resentment, p 174.  
Strawson here is concerned with the person/ soul relation, whereas I have utilised the spirit of the argument to the analogous case of the bodied being/person relation.
7. Cf Kan, Critique of Pure Reason (ch on the paralogisms) A363-4,

footnote. Kant here was dealing with the hypothesis that it is the sameness of the soul substance or Pure Ego which constitutes the sameness of the person. His famous retort to this hypothesis is; "whenever you say there is one continuous soul-substance, I say there is a whole series of them each of which transmits its states, and the consciousness of them, to its successor, as motion might be transmitted from one to another of a whole series of elastic balls". ) (Strawson's paraphrase of Kant, Ibid, p 233)

8. JP, 1962, p 403.
9. See Lewis and Flew, symposium on "Survival" PAS (suppl) 1974-75; and also G. Englebretsen, Speaking of Persons (Dalhousie Univ. Press, Halifax, Nova Scotia, 1975), Ch 3. See also F. Cowley, "The Identity of a Person and His Body", JP 1971.
10. "Self, Mind and Body", Freedom and Resentment.
11. See Copleston, Aquinas (Pelican original, Penguin 1955), p 170 and Strawson, Individuals, p 103.
12. See Copleston, Ibid, Ch 4 esp pp 163 and 168. CF Aquinas, Summa Theologica, Ia, 89, 1.
13. See Copleston Ibid, esp pp 160 and 167.
14. J. Harrison "Embodiment of Mind .. ", PAS 1974. Penelhum also concedes such possibilities presumably to make the idea of disembodied existence plausible enough; Survival and Disembodied Existence, chs 2 and 3.

#### Chapter 4: Section (iv).

1. This we have made clear in the first chapter (section ii). Arguments to show that knowledge of self-identity (as distinct from personal identity) not only can be, but is, non-criterial have been adduced by Shoemaker (especially, Self-Knowledge and Self-Identity)

and also by Strawson (Bounds of Sense, pp 164-5).

2. The suggestion, in view of the essential imperceptibility of the disembodied persons, that the latter has to live in a solipistic world has appealed to many as irresistible (see, for example, Penelhum, Survival and Disembodied Existence): and some even have wanted to draw this conclusion on conceptual grounds (Cf also Penelhum, Ibid; this is also implicit in Shoemaker's book referred to above, especially in his attempt therein to show that the relation between a person's seeing and his having the 'point of view' on his body is not contingent, id, Ch 4 secs 3-6). Somewhat similarly implicit arguments can also be found in Puccetti's Persons (Macmillan, London 1968, Ch 1). However, as I do not intend to argue that the idea of disembodied survival is logically impossible or straightforwardly nonsense, I will leave the last word on this to the proponents of the thesis and allow that the disembodied persons could be said to make memory claims.
3. See pp 256-7 and 217 above.
4. PAS 1974, pp 48-49 (my emphasis).
5. op cit, p 49 (few lines later).
6. This point became explicit in course of Malcolm's discussion of his paper mentioned above.
7. See, for example, Penelhum, Survival and Disembodied Existence, esp p 77. Though, on my showing, Penelhum would be right in saying this about "non-embodied" persons.
8. Indeed, this fact of "inheriting" the understanding of the P-predicates has been thought to be an essential requirement of the disembodied persons being able to make statements containing P-predicates or even for our making such statement about them. (See, for example, Penelhum, Ibid, p 22)

9. While the Roman Catholic tradition generally believes that the pre-mortem body and the resurrection body are one and the same, Aquinas, in particular, lays special emphasis on their being the same so much so that, according to him" ... if it be not the same body it will not be a resurrection, but rather the assuming of a new body" (Summa Tehologica, III Q 79 A1). Also cf Ibid, Q79 A2, where it is claimed that resurrection will be effected only by "the self-same soul being united with the self-same body", since, presumably, (as Aquinas did belive, see last section) the soul by itself cannot constitute a human person but can do so only in its union with a body. (So that if  $S_1 + B_1 = P_1$ , taking 's for a soul, B for a body and P for a person, then even if  $S_1 + B_2$  may make up a person, it cannot make the same person). But here again, except for a pious hope that the soul will rise together with the same body ( i.e. perhaps, God will recreate or reconstitute the resurrection body out of the persisting spatio-temporally continuous elements of the original pre-mortem body), no reasonable guarantee is provided so that this body will not be just an exactly similar body rather than the same body.
10. "Theology and Verification", especially in his third picture of instantaneous change of bodies.
11. Essay, II Ch 27 sec 1.
12. "Survival, Reincarnation and the Problem of Personal Identity", Journal of the Indian Philosophical Association, July-December, 1968.
13. op cit, p 138.
14. op cit, p 139.
15. Self-Knowledge and Self-Identity, pp 239-41, (cf also Wittgenstein PI pt II p 56 bottom line).

16. Cf PI, pt II p 56.
  17. Ibid, pt I sec 142.
  18. Survival and Disembodied Existence, see chs 9 and 10, and esp p97.
  19. Ibid, p 95. Also see his "Personal Identity, Memory and Survival", JP 1959 pp 900-901. Penelhum's account of such a person and personal identity should be entertained only as a reductio as, in fact, he intended to use this. But one enigmatic feature of his analysis is that he seems to allow that persons could (particularly in the Hick-type cases referred to in the text earlier) be coherently and intelligibly construed as gap-inclusive entities in the described way. For an unqualified commitment to this will not only be unacceptable to us, but will expose an inherent inconsistency in his own position. He cannot, i.e., have it both ways: that bodily continuity is necessary and sufficient condition of personal identity (see Id, p 67 ) and that persons could be thus construed.
- Our qualification with regard to this will be, as it has been with regard to disembodied survival, that though our talk in such cases of "same person" is intelligible, it succeeds in being so only in a secondary sense.
20. See ch 2 sec iii, where it has been argued that, on a purely "stream of consciousness" account, no guarantee would be there that a "soul-phase" may not fall into the wrong stream.

Chapter 5: section (1).

1. Cf Wittgenstein: "We are not forced to talk of a double personality (by which he means "two persons", PKM)", The Blue Book, p 62.
2. See J.M. Shorter, "Personal Identity, Personal Relations and Criteria", PAS 1970-71, and esp Wittgenstein, The Blue Book, pp 61 and 62. It is also explicit in J.F.M. Hunter, "Personal Identity" in Essays After Wittgenstein, esp p 41.

3. E.g. Wiggins, ISTC, pp 52 and 53.
4. This view of Parfit has been fully discussed and dismissed in ch 3 sec iv (and also in Appendix 2). For an explicit exposition of this interpretation of Parfit see Swinburne, "Personal Identity", PAS 1974, p 243 and note 9 to that paper.
5. Cf Wiggins ISTC, pp 52-53.
6. To borrow a fitting example from Wiggins, cf Ibid, p 55.
7. As Shoemaker, for example, had wanted to say in his famous Brownson case (Self-Knowledge and Self-Identity, pp 30 and 246). However, if our modification in the concept of bodily continuity (see ch 3 secs ii and iii) is justified, Shoemaker's plea will amount only to a changed or modified version of our normal criterion (i.e., bodily continuity) - and not to a change or modification of the criterion itself.
8. Penelhum, Survival and Disembodied Existence, p 83.
9. Cf Ibid, p 85. We have explicated the unclarity of these situations by showing them to be worthy, at best, of a secondary, degenerate description - besides being dubious.

Chapter 5: section (ii).

1. In the sense in which they have been analysed in earlier chapters, especially ch 1.
2. The weakness of such a theory, in so far as it purports to lead to an entirely non-physical account of personhood and personal identity, has been discussed at some length in chapter 1 sec ii and ch 2 sec ii. In view of the extreme familiarity of this weakness, I do not intend to defend the "Simple View" in this obvious formulation - although I do defend another formulation of the theory which I think is correct.
3. See Parfit, "Personal Identity" (PR 1971) and "On 'The Importance of

Self-Identity" "(JP 1971); and see also Lock4 Essay, II Ch 27 sec 18.

4. By contrast, there are certain other words which have a non-unitary role; for example, 'intelligence', 'strength', 'reasonableness' and the like. It normally makes no sense to talk of a unit of intelligence or even of strength (except on the technical sense in which a minute is called a unit of time), and accordingly people may be quite sensibly be said to be more or less intelligent, strong etc. It may at times be more sensible to regard a stronger man as equal to two or more men who are less strong; but the former is still as much a person as any of the latter two
5. As Parfit, in particular, has suggested. (See Ch 4 sec v above; see also note 6 below).
6. Parfit PR 1971, JP 1971 and "Later Selves ..." (in Philosophy and Personal Relations, ed Montefiore).
7. Note his emphatic claim that "... in its nature - in what it involves - personal identity is a matter of degree".
8. For yet another recent attack on this dubious separation, see Brian J Smart, "Diachronous and Synchronous Selves" (see iv), in the Canadian Journal of Philosophy (CJP , hereafter) March 1976.
9. See Ch 3 sec v for fuller arguments against this.
10. See especially his "Later Selves ..." in Montefiore.
11. As Geach has rightly remarked about the Lockean theory of personal identity see his God and the Soul, p 4.
12. As Parfit rightly pointed out, see his "Later Selves ...", Montefiore p 157.
13. See PR 1971, p 25 and Montefiore, p 141.

14. See Montefiore, p 141.
15. See CJP March 1976.
16. "Later Selves ... " Montefiore, p 111.
17. See, for example, Wiggins, "On Being in the Same Place at the Same Time", PR 1968, p 91.
18. CJP March 1976, pp 24-25.
19. See Williams "The Self and the Future", and also his "Are Persons Bodies ?". See also Swinburne ("Personal Identity" PAS 74 p 244) for the importance of the survivor being me.
20. The Blue Book, p 25.
21. Cf JFM. Hunter "Personal Identity" in Essays after Wittgenstein, pp 41-42.

Appendix - 1.

1. See Strawson, The Bounds of Sense pp 164-67, for more on this. The argument that the belief in criterion-less self-knowledge does not imply the non-physical nature of the subject of experience or of self-knowledge (i.e. person) is fairly implicit in his Individuals, Ch 3 sec 6.
2. Cf Strawson, Individuals, p 111.

Appendix - 2.

1. This suggestion is expressly made by Wiggins (ISTC, p 55, para 2) and, as I take it, implicitly by Parfit in the doctrine that "non-branching psychological continuity" can guarantee personal identity. (Cf "Personal Identity", PR 1971).

2. I owe the formulation of this suggestion directly to Professor Swinburne. He himself is strongly in favour of this suggestion as is also implicit in his theory of unanalysability of personal identity (See his "Personal Identity" PAS 1974, esp p 241).
3. See Swinburne op cit, p 243.  
For our argument to the same effect see Ch 3 sec v above.
4. Those who would think that the result of the operation may be otherwise and that the two halves of the brain may be heterogeneous in the relevant respects, and so may be the resulting persons, would not be making any mistake; they would rather be pointing to what actually is the case rather than what might be possible. But in that case, the problem of whether one or each "product" is the same person as the original person will not trouble us any more. For in that case our argument would take a different line to show that none of the resulting persons can be identical with the original person, by casting doubts on the proposition that any of them is a person in the first place. This can be argued by drawing attention to what Thomas Nagel calls the "functional duality of the cerebral cortex" (See his "Brain Bisection and the Unity of Consciousness", in Synthese 1972) together with the unitary concept of a person. On the basis of his findings, Nagel's own conclusion is that there would be no whole number of individual minds that the patients (of bisected brain transplants) can be said to have (See op cit, pp 402 and 409), though they could still be said to engage in mental activities (op cit, p 402). My further conclusion from this would be that if the idea of any (countable number of ) individual minds goes, the idea of person or persons

goes too. The idea behind the last conclusion is that unless all that is necessary and sufficient for the concept of a person is transferred, then the idea of a person breaks down; and on the assumption that the two halves are heterogeneous ( that the two halves must co-operate in the controlling of the personal capacities) we will not be transferring all that is necessary and sufficient for the concept of a person.