

T H E ' N A T U R E O F A S E L F

by

G. le Chat

A thesis submitted to Rhodes University, in
fulfilment of the requirements for the degree
of Master of Arts in Philosophy.

January, 1978

ACKNOWLEDGEMENTS

Firstly, I want to thank my supervisor, Dr. James Moulder, for his valuable comments and criticisms of my work, and for his constant encouragement during my writing of this thesis. Secondly, I want to thank those of my colleagues at the University of Cape Town with whom I discussed certain chapters of this thesis. I benefitted considerably from the discussions I had with them. Thirdly, I must thank the Human Sciences Research Council for their financial assistance. The arguments and the conclusions reached in this thesis, however, are my own and are not intended to reflect those of the Human Sciences Research Council.

CONTENTS

<u>Chapter</u>	<u>Page</u>
INTRODUCTION	1
1 THE NO-SUBJECT DOCTRINE OF THE SELF	5
(a) Introduction	5
(b) Wittgenstein, Hume and Russell : exponents of the no-subject doctrine	6
(c) Evaluation of Wittgenstein's argument	10
(d) Evaluation of the bundle theorists' argument	15
(e) Summary	24
2 STRAWSONIAN DUALISM	26
(a) Introduction	26
(b) Is Strawson neutral between dualism and materialism?	27
(c) Can Strawson offer a satisfactory explanation of the notion of 'a logically adequate criterion?'	35
(d) Summary	46
3 CARTESIAN DUALISM	48
(a) Introduction	48
(b) Strawson's objection to Cartesianism	49
(c) Lewis' defence of Cartesianism	55
(d) A reply to Strawson's objection	63
(e) Good reasons for rejecting Cartesianism	68
(f) Summary	74
4 MATERIALISM AND PERSONAL IDENTITY	76
(a) Introduction	76
(b) Strawson and Locke on Personal Identity	76
(c) Memory as a criterion of personal identity	79
(d) A prejudice in favour of bodily identity	86
(e) Identity and Survival	92
(f) Williams' Puzzle	96
(g) Summary	103
CONCLUSION	104
Bibliography	110 - 113

INTRODUCTION

1.1 The purpose of this thesis is to demonstrate just what kind of entity a self or person is. One of the most popular answers to the question 'What is a self or person?' or 'What am I?' has been the one provided by Descartes. He says in the Sixth Meditation

' . . . from the mere fact that I know with certainty that I exist, and that I do not observe that any other thing belongs necessarily to my nature or essence except that I am a thinking thing, I rightly conclude that my essence consists in this alone, that I am a thinking thing, or a substance whose whole essence or nature consists in thinking.'¹

By 'a thinking thing' or 'a substance' he means something which is 'entirely and truly distinct from my body' and which 'may exist without it'. Descartes' view of a person as an essentially incorporeal entity has, in recent times, been severely criticized, but it is, as we shall see, still thought by some to be defensible.

1.2 A more recent, and perhaps more appealing, answer to the question 'What is a self or person?' has been provided by P.F. Strawson. A person, he maintains is 'a type of entity such that both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics . . . are equally applicable to a single individual of that single type.'²

¹Descartes, R., Discourse on Method and the Meditations trans. F. E. Sutcliffe (Penguin Books), p 156.

²Individuals (Methuen, London, 1959), Ch.3.

The concept of a person is not, however, to be analysed, he says, as an entity consisting of a particular consciousness and a particular human body. For the concept of a person is 'logically primitive' or 'logically prior to that of an individual consciousness.'¹ Strawson enlarges on his initial statement of what a person is by claiming that two kinds of predicates are properly applied to persons. These are M-predicates and P-predicates. M-predicates are predicates which are also properly applied to material bodies. They include things like 'weighs 10 stone', 'is in the drawing room', and so on. P-predicates, on the other hand, are predicates which we would not dream of ascribing to material bodies; they are applied exclusively to persons. There are, it seems two classes of P-predicates. Strawson distinguishes, for example, such predicates as 'is smiling', 'is going for a walk', 'believes in God', and so on.² The latter class of P-predicates he calls 'predicates ascribing states of consciousness'. A common characteristic of all the P-predicates is that they all 'imply the possession of consciousness on the part of that to which they are ascribed.'³

1.3 Now both Descartes and Strawson, it can be seen, think that a person is an entity which has psychological attributes (though Strawson appears to think that the entity which has these attributes is essentially a material entity). There are some philosophers, however, who appear to have challenged the truth of the assertion that a person is an entity which has

¹ op.cit., p. 103.

² op.cit., p. 104.

³ op.cit., p. 105.

psychological or mental attributes. Strawson refers to the view of a self held by those philosophers who challenge the truth of this assertion as the no-ownership or no-subject doctrine of the self. 'On this view', says Strawson, 'it is only a linguistic illusion that one ascribes one's states of consciousness at all, that there is any proper subject of these apparent ascriptions, that states of consciousness belong to, or are states of, anything.'¹ I shall examine the no-ownership view in the first chapter of this thesis. The main aims of the chapter will be to show (i) that the arguments which have been advanced for the rejection of the view that a self is a subject of experiences are seriously defective;² and (ii) that a self must be an entity to which experiences belong.

1.4 Having rejected the no-ownership or no-subject view of a self I shall, in the second chapter, examine Strawson's view of a person, i.e., the view that a person is an entity to which both material body predicates and psychological predicates are equally applicable. One of the main difficulties with Strawson's thesis, it will be seen, is ascertaining exactly what his conception of a person is. For although his claim that a person is 'a type of entity such that both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics . . . are equally applicable to a single individual of that single type 'is a'topic neutral claim.'³ and thus compatible with the view that the concept of a person is primitive or unanalysable, it appears questionable whether he

¹ op. cit., p.94.

² By the term 'a subject of experience' I mean nothing more than an entity to which psychological or mental attributes belong.

³ This expression belongs to J.J.C.Smart. See his 'Sensations and Brain processes' *Philosophical Review*, 68, 1959.

is neutral between the dualists' position and the materialists' position. I shall argue that Strawson's conception of a person is, in fact, very similar to Descartes' (in spite of the fact that he attempts to refute Descartes' view).

1.5 Having discussed Strawson's conception of a person I shall then go on to examine Descartes' view, i.e., the view that a person is essentially an immaterial entity. The main aims of the third chapter will be : (i) to present Strawson's argument for the rejection of Cartesianism as clearly as possible; (ii) to show that H.D. Lewis' attempt to meet Strawson's objection fails; (iii) to show that it is not impossible for a defender of Cartesianism to meet Strawson's objection; and (iv) to show that even though Strawson's objection does not work, there are still good reasons for rejecting the view of a person as an 'immaterial substance.'

1.6 Finally, in Chapter five, I shall attempt to provide more detailed support for the view that a person is a material thing, by considering whether the problem of personal identity is to be decided in favour of the bodily criterion or the memory criterion. I hope to reveal that the bodily criterion is the more fundamental of the two. In this way, then, I hope to provide support for the materialist view of a person.

CHAPTER I

THE NO-SUBJECT DOCTRINE OF THE SELF

Introduction

1.1 Most of us regard the assertion that a person is an entity which has psychological or mental attributes as an obvious truism, but the truth of this assertion, as I have already said, appears to have been challenged by some. Strawson maintains that he is not sure whether anyone has explicitly held the no-ownership or no-subject view of the self, but that 'there is some evidence that it was held, at one period, by Wittgenstein and possibly also by Schlick.' I do not intend to evaluate his discussion of the no-ownership theory, mainly because his discussion is so unclear and confusing,¹ but I do wish to point out, first, that Wittgenstein and Schlick are not the only philosophers who have seemed to deny that a person is an entity which has psychological attributes, that Hume and Russell, for example, have also seemed to deny this; and, second, that Wittgenstein did not, as Strawson suggests, just hold the theory for a certain length of time and then abandon it. Having done this I shall evaluate Wittgenstein's, Hume's and Russell's arguments for the rejection of the view that persons are entities which have psychological attributes, or, in other words, that persons are subjects of experience. In evaluating Wittgenstein's argument I hope to show that although

¹ The reason that his discussion is confusing is, I believe, due largely to the fact that Strawson is not so much concerned with demonstrating that there must be subjects of experience, perhaps because he thinks that this is so obvious, but with establishing the nature of subjects of experience. In other words, he is not so much concerned with the question 'Is the self a subject?' but with the question 'What is the nature of the self?' or 'Is the self a material or immaterial subject?'

his argument can easily seem highly plausible, largely in view of the fact that the word 'I' does at first glance appear to function more like the word 'here' than a proper name, it fails because it does not take account of all the facts. In evaluating Hume's and Russell's very similar arguments I hope to show (i) that if their assumption that we observe our mental states was correct then it would force them to concede that a self is a subject of experiences; (ii) that their assumption is, in fact, mistaken; and (iii) that even though it is a mistake to think of ourselves as subjects observing our mental states, the Humean view of a person is still unacceptable.

Wittgenstein, Hume and Russell : exponents of the no-subject doctrine

2.1 Strawson maintains that "the evidence that Wittgenstein at one time held the no-ownership or no-subject doctrine of the self" is to be found in Moore's articles in Mind on 'Wittgenstein's Lectures in 1930-33' (Mind, Vol. LXIV, pp 13-14).¹ This suggests that Wittgenstein held the view for a time and then abandoned it. The fact of the matter, however, is that he never did abandon the view. Both in The Blue Book, dictated during 1933-34, and Philosophical Investigations, written in the 1940's, he rejects the view that the word 'I' in sentences like 'I have a toothache' and 'I think it will rain' refers to a possessor of experiences.¹ In The Blue Book Wittgenstein writes,

"Now the idea that the real I lives in my body is connected with the peculiar grammar of the word 'I', and the misunderstandings this grammar is liable to give rise to. There are two different cases in the use of the word 'I' (or 'my') which I might call 'the use as object' and 'the use as subject'. Examples of the first kind of use are these : 'My arm is broken', 'I have grown six inches', 'I have a bump on my forehead', 'The wind blows my hair about',
Examples of the

¹ See: The Blue and Brown Books (Basil Blackwell, Oxford, 1958), pp. 64-74, and Philosophical Investigations (Basil Blackwell, Oxford, 1953) pp. 120-125.

second kind are: 'I see so-and-so', 'I hear so-and-so', 'I try to lift my arm', 'I think it will rain', 'I have a toothache' " (pp 66-67)

Wittgenstein explains the difference between cases in which the word 'I' is used 'as object' and cases in which it is used 'as subject' by maintaining that the cases in which it is used 'as object' 'involve the recognition of a particular person, and there is in these cases the possibility of an error'. One can be mistaken, according to Wittgenstein, in thinking that one's arm is broken, and that one has a bump on one's forehead. 'It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbour's. And I could, looking into a mirror, mistake a bump on his forehead for one on mine'. But one cannot, he says, be mistaken in this way when one says 'I have a toothache' or 'I am in pain', i.e., one cannot mistake someone else who has a toothache or who is in pain for oneself. Wittgenstein concludes that to say 'I am in pain' is thus more like moaning than a statement about a particular person. He goes on to point out that there is a tendency to think that what the use of the word 'I' in a sentence like 'I am in pain' does not involve is the identifying of a particular person by his bodily characteristics, but this, he says, 'creates the illusion that we use this word to refer to something bodiless, which, however, has its seat in our body.'¹ It is an illusion to think that the word 'I' refers to an ego (or anything else) because there is, in Wittgenstein's view, no question of identifying anything when one says 'I am in pain'.

2.2 That Wittgenstein continued to hold the view that the word 'I' does not denote a possessor of experiences when he came to write the Investigations is abundantly clear. He says,

¹ op. cit., p.69.

for example,

"When I say 'I am in pain', I do not point to a person who is in pain, since in a certain sense I have no idea who is." And this can be given a justification. For the main point is I did not say that such-and-such a person was in pain, but "I am . . ." Now in saying this I don't name any person. Just as I don't name anyone when I groan with pain. Though someone else sees who is in pain from the groaning.¹

In support of the claim that the word 'I' does not name a person Wittgenstein says, as he does in The Blue Book, that one cannot doubt whether it is oneself or someone else who is having a particular experience. He then maintains that the word 'I' functions more like the words 'here' and 'this'. He says "I' is not the name of a person, nor 'here' of a place, and 'this' is not a name. But they are connected with names. Names are explained by means of them."²

2.3 But enough of Wittgenstein for now. Other philosophers who have seemed to deny that a person is a subject of experience have, in my opinion, often held one version or another of the 'bundle theory' of the self. This theory is also sometimes called the 'serial theory' and the 'logical construction theory', and has been held by a number of twentieth century philosophers, including Russell. The originator of the 'bundle theory', however, was Hume. Hume denied in the Treatise that we can have an idea of 'self or person' or 'self or substance' as 'something simple and continued' and concluded that a person is 'nothing but a bundle or collection of different perceptions'.³ He expanded on this statement of what a person is when he remarked

¹ op. cit., p 122, paragraph 404.

² op.cit., p 123, paragraph 410.

³ Hume, D., A Treatise of Human Nature, ed. L.A.Selby-Bigge (Oxford, 1888), Book 1, Part 4, Sec. 6.

that

"The mind is a kind of theatre, where several perceptions make their appearance; pass, repass, glide away and mingle in an infinite variety of postures and situations."

In short, although he shared the Cartesian assumption that a person is essentially a mind, he rejected the idea of the mind as a simple persisting entity to which experiences belong. The existence of a particular 'perception' ('perception' is Hume's most general term for a mental event) does not, therefore, involve, in Hume's view, the existence of anything that has the perception. To say that a particular person has a pain is not to say, as a Cartesian thinks, that a pain belongs to, or is owned by, a certain simply persisting entity, but is to say that a pain is a member of a certain 'collection' of perceptions. Was Hume, then, in rejecting the notion of a Cartesian subject of experiences rejecting the view that a person is an entity to which psychological attributes belong? I am inclined to think that he was, for it a person is identical with his mind, and the mind consists of nothing over and above its various states, it seems to follow that a person cannot be something that has these states.

2.4 Let us now turn to Russell. Russell advances his version of the 'bundle theory' in The Analysis of Mind. He says, "It is supposed that thoughts cannot just come and go, but need a person to think them. Now, of course it is true that thoughts can be collected into bundles, so that one bundle is my thoughts, another is your thoughts and a third is the thoughts of Mr Jones. But I think that the person is not an ingredient in the single

thought : he is rather constituted by relations of the thoughts to each other and to the body . . . The grammatical form 'I think', 'you think', and 'Mr Jones thinks', are misleading if regarded as indicating an analysis of a single thought. It would be better to say 'it thinks in me', like 'it rains here'; or better still, 'there is a thought in me!'" ¹

Russell seems here to be saying, like Hume, that it is a mistake to think of a person as an entity to which thoughts belong, for a person (or at least the mind of a person) is simply a collection of thoughts. Now it might be suggested that what Russell meant to say is simply that it is a mistake to think of a person as a Cartesian subject. But it is not at all clear that he did mean only this. For if a person is, in some sense, a subject of experiences then why should the grammatical form 'I think' be misleading? Why should it be better to say 'it thinks in me' like 'it rains here'? If it is more accurate to say 'it thinks in me' it must be so, it seems, because the word 'I' implies that there is a subject.

Evaluation of Wittgenstein's argument

3.1 Wittgenstein's argument for the rejection of the view that the word 'I' refers to a possessor of experiences may, briefly, be summarised as follows. When someone says 'I am in pain' or 'I am depressed' there is no question of his having to identify anything. This is supported by the fact that one cannot doubt whether it is oneself or someone else who is in a particular mental state. Only if saying something like 'I am

¹ Russell, B., The Analysis of Mind (Allen & Unwin, London, 1921), pp. 17-18.

in pain' involved the identification of a particular entity would there be a possibility of error. The statement 'I am in pain' is, thus, no more a statement about a particular person than moaning is, and the 'I' in this statement functions more like the word 'here' than a proper name. This argument can, I believe, easily seem highly plausible. Firstly, Wittgenstein is clearly correct in claiming that there is no question of one's having to identify anything when one says something like 'I am in pain', and that one cannot mistake someone else's pain for one's own. And, secondly, the word 'I' can, I believe, easily seem to function more like the word 'here' than a proper name. Godfrey Vesey, in his Personal Identity, has, in fact, attempted to support the view that the word 'I' functions like the word 'here'.

3.2 Vesey, in the second chapter of Personal Identity attempts to show (i) that the word 'here' can be used meaningfully without our knowing something which we call 'here', and (ii) that the word 'I' is like the word 'here' in this respect. In his attempt to demonstrate (i) he considers the case where his wife calls out 'Where are you?' and he replies 'Here'. His wife knows, from this, says Vesey, roughly where he is, not because he has said where he is, but because she can hear from what direction his voice is coming. Vesey points out that even if he did not know where he was he would be entitled, on hearing his wife, to say 'I'm here'. If he had been knocked unconscious by a burglar, blindfolded, and put into a cupboard, he could, on regaining consciousness and hearing his wife, call out 'I'm here'.

without being accused of "using the word 'here' without meaning".¹ For the word 'here', Vesey explains, can have a use without there being some thing in the world (in this case, a place) which the word stands for. "People know where I am, on hearing me say 'Here', not in virtue of knowing what I mean, but in virtue of being able to locate sounds."² Vesey attempts to strengthen his argument by drawing two distinctions, viz., between stating where someone is and indicating where someone is, and between something having, and not having, a truth value. With regard to the first distinction he says that "if I say 'I'm in the study' I am stating my whereabouts (though not who 'I' am). But if I say 'I'm here' I am merely indicating my whereabouts."³ He uses the second distinction to support this, arguing that the utterance 'I'm here' is not a statement about his whereabouts because it is not an utterance which can properly be said to be right or wrong.

3.3 But let us now leave Vesey's discussion of the word 'here' and turn to his discussion of the word 'I'. He attempts to show that the words 'I' and 'here' have a similar function by comparing the use of 'I' in soliloquy to the use of 'here' in soliloquy. On the use of 'here' in soliloquy Vesey has the following to say.

"Suppose my wife and I are lost, in cloud, on a mountain. We try to keep in touch by shouting occasionally 'Where are you?' 'Here'. But we drift apart, and can no longer hear one another. Still lost, I say to myself, 'Where am I?' and reply 'Here'. But whereas my wife could have learnt in what direction I was from her had she heard me, there is nothing I can learn from it.

¹ op.cit., p.24.

² op.cit., p.25.

³ ibid.

It is, as Geach would say, 'idle, superfluous'."¹

The word 'I' in soliloquy, in Vesey's view, is also 'idle', though he thinks it too could convey information to others if they heard him talking to himself. He advances the following example in support of this.

" Suppose I have lost my memory. All I can find in my pockets is a scrap of paper with a telephone number written on it. Perhaps if I ring the number the person who answers will know who I am . . . I am all set to say 'It's me', but nobody answers the phone. Emptily I say to myself: 'It's me, whoever I am'."²

The word 'I', then, in Vesey's view, is like the word 'here' in that it is also 'an indicating word', i.e., others can learn something from it, e.g., who the speaker is.

3.4 Now Vesey and Wittgenstein are correct, in my opinion, in holding that the word 'I', like the word 'here', is, from the speakers point of view, 'idle'. But it does not follow from this that the 'I' cannot be said to refer to anything. To draw this conclusion is to fail to realise that we say such things as 'I am in pain' and 'I am depressed' primarily for the sake of others. When one wants to inform another that some other person is, say, feeling depressed, one either says 'he is depressed' or 'So-and-so is depressed', and the words 'he' and 'So-and-so' obviously refer to a particular person. But when the one who is depressed is oneself one does not say 'he is depressed' or 'So-and-so is depressed', and this is not because one is not referring to a particular person. The reason that one does not say 'he is depressed' or 'So-and-so is depressed'

¹ op.cit., p.30.

² op.cit., pp.30-31.

is because one wants to inform others that it is oneself who is depressed, and one wants to do so in such a way that there is no uncertainty about who is depressed, i.e., one wants to make it quite clear that it is oneself who is depressed. One could, perhaps, say 'So-and-so is depressed' when the person who is depressed is oneself - this would be somewhat unusual, though not absurd - but, in order to avoid confusion, it is better to say 'I am depressed', for one may be in a situation in which there are a number of people with the same name as oneself. We can, therefore, agree with Wittgenstein that a person who says 'I feel depressed' does not have to employ any criteria in order to make this claim, and that he cannot mistake someone else's depression for his own, but this does not prove that he is not using the 'I' to refer to himself, that the 'I' is superfluous, for he uses the 'I' to refer to himself for the sake of others.

3.5 One final point: Both Vesey and Wittgenstein, we have seen, maintain that although the word 'I' in a sentence like 'I am in pain' does not refer to a person it does indicate who is in pain. But it is not clear what is to be gained by saying this. For does not the claim that the 'I' indicates who is in pain also imply the existence of an entity to which experiences belong? It might be thought that the 'I' could be said to indicate who is in a particular mental state, without presupposing the existence of a subject, if persons were bundles of experiences. If persons are bundles of experiences, it might be said, then the word 'I' cannot be said to refer to a subject of experiences, and merely indicates that the mind which

is connected with the body from which the words 'I am in pain' are uttered is in a particular mental state. I hope to show in the following section, however, that the Humean view of a person is open to serious objections.

Evaluation of the bundle theorists' argument

4.1 The reason why Russell rejects the view that a self is a subject is because, in his own words, 'the act in thinking is not empirically discoverable, or logically deducible from what one can observe.'¹ Hume, too, may be interpreted as denying that a self is a subject on the ground that no entity to which experiences belong is revealed in introspection.² For he says "For my part, when I enter most intimately into what I call myself I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe anything but the perception."

Ayer also rejects the view that a self is a subject on the ground that the substantival ego is 'an entirely unobservable entity'. He says 'It may be suggested that it is revealed in self-consciousness but this is not the case . . . But if the substantive ego is not revealed in self-consciousness, it is not revealed anywhere, the existence of such an entity is

¹op.cit., p.18.

²But it should be noted that although he may be interpreted as rejecting the view that a self is an entity to which experiences belong on the ground that no such entity is observable in introspection, he rejects this view, not so much because he cannot observe an entity to which experiences belong but rather because of his belief that a thing does not consist of something over and above its perceived qualities. In other words, he criticizes Descartes' conception of the mind for the same reasons that Berkeley criticized Locke's theory of substance. Locke appears to have held, like Descartes, that material objects do not consist solely of their perceived qualities or properties, that there must be a "substance" which supports the qualities or properties of a material object. He denied, however, that we can have a clear and distinct idea of a 'substance' which has qualities or properties. Berkeley criticized this view of substance on the ground that we can talk meaningfully only about what we are acquainted with.

completely unverifiable.¹ In short, the reason why the bundle theorists have denied that a self is a subject is because, when they have introspected, they have been unable to observe any subject to which experiences belong.

4.2 There is, however, a serious difficulty with denying the view that a self is an entity to which experiences belong because no such entity is 'empirically discoverable'. What Hume and the other bundle theorists fail to realise is that in regarding introspection as analogous to perception, they are forced to concede that a self is a subject of experiences. If introspection involves the observation of one's mental states then there must be something which observes these states, for nothing can be observed unless something exists to observe it. Observation requires both a subject and an object. The reason that Hume fails to find a subject, it might be said, is simply because when a perceiver perceives an object he is automatically excluded from what he is perceiving. As Shoemaker remarks : "One cannot see one's eyes (except in a mirror) if one is seeing with them; the place from which one sees is necessarily excluded from one's field of vision."²

4.3 Locke and Berkeley, it should be noted, also held that we observe mental states, that mental states are the objects of 'inner sense', though, having more insight than Hume, they did not conclude that because only mental states are revealed in introspection that a self consists only of a 'bundle' of experiences. They both held that we acquire such ideas as perceiving, thinking, doubting, knowing, etc., by observing the 'internal operations' of the mind. Locke, for example, remarked that "the mind, in

¹ Ayer, A.J., Language Truth and Logic, p.166.

² Shoemaker, S. Self-Knowledge and Self-Identity, Ch. 2, p.78.

all its thoughts and reasonings, hath no other immediate object but its own ideas, which it alone does or can contemplate."¹ Berkeley remarked, rather similarly, that "it is evident to any one who takes a survey of the objects of human knowledge, that they are either ideas actually imprinted on the senses; or else such as are perceived by attending to the passions and operations of the mind."² But perhaps Locke, Berkeley and Hume are all mistaken in their assumption that we observe our mental states. It is, I think, worth digressing for a moment to examine this assumption more closely, for if it is true then it would seem that there is a good ground for holding that a 'self or person' is a subject of experience.

4.4 Shoemaker provides an excellent discussion of the question of whether a person observes his mental states or not in his Self-Knowledge and Self-Identity. He begins by suggesting that the reason why it is held that we have knowledge of our mental states on the basis of observation is because it is thought that every contingent fact a person knows he either knows directly on the basis of observation or inferentially on the basis of observation.³ He then points out that if knowing that he is in a particular psychological state involves some sort of observation it must involve that he observes something. It must involve observing some mental entity like a pain or a thought or an image, etc. But, says Shoemaker, nothing can be observed unless something observes it.⁴ So if he knows that

¹Locke, J., Essay Concerning Human Understanding, ed. A.C. Fraser (Oxford, 1894), Bk. IV, Ch. 1, Sec. 1.

²Berkeley, G. The Principles of Human Knowledge, Part 1, Sec. 1; Works of Berkeley, ed. G. Sampson (London, 1897) Vol. 1, p. 181.

³op.cit., p. 63.

⁴It is the point, we have seen, which Hume overlooks.

he is in pain on the basis of observation there must be a subject that perceives the pain. And since it is he that knows that he is in pain, it must be he that observes the pain. It is in this way, Shoemaker believes, that the view that we observe our mental states forces upon us the picture of a person as a subject perceiving various kinds of mental objects. Shoemaker eventually maintains, however, that the view that we observe our mental states is a mistaken one.¹ He thinks that it is senseless to claim that we perceive or observe mental phenomena like pains, thoughts, images, etc. Our so-called 'awareness' of our mental states cannot, in his opinion, be regarded as like observation at all. Let us now consider the argument he uses to substantiate this view.

4.5 Shoemaker argues that what justifies a person making a 'first-person psychological statement' such as 'I am in pain' is simply his being in pain, not his having evidence of being in pain or his observing the pain.² There is, according to him, a strong philosophical inclination to explain what is involved in saying 'I am in pain' as follows. When I assert that I am in pain this does not involve any kind of inference. I do not infer the fact that I am in pain from some other fact. When I assert that I am in pain what I am directly acquainted with is the pain itself. This direct acquaintance with the pain itself justifies me in saying 'I am in pain'. Thus if by 'criterion for saying that I am in pain' is meant something that shows that I am in pain then there is a criterion, viz., the pain itself. And this criterion is a private one for no one else can be shown that I am in pain in this way. But this

¹op. cit., Ch.6, pp. 211-224.

²op. cit., p. 217.

explanation, says Shoemaker, 'rests on a mistake'. It rests on the false belief that 'acquaintance with' or 'awareness of' mental objects is a kind of perception or observation. The statement 'I know I am in pain because I am aware of a pain' cannot, he maintains, be informative or explanatory in the way in which the statement 'I know there is a tree on the hill because I see one there' can. The latter statement implies that there is more than one way of knowing that there is a tree on the hill, but it is not true of the former statement that it implies that there is more than one way of knowing that one is in pain. And if someone were to assert that he knows there is a tree on the hill because he sees one there it would be possible to check whether he is justified in making this assertion - by determining, for example, whether his eyes are open, and whether his eyes are directed towards a tree, etc. - but it is not true that we can check whether someone is justified in asserting that he is in pain.

4.6 Shoemaker thus concludes that if being aware of a mental state like pain involved observing a pain then it would have to be the case, (i) that it can be an open question, to be settled empirically, whether a person who is in pain and thinks he is in pain is in fact aware of pain; (ii) that it is possible for a person to be in pain without being aware of a pain; (iii) that being aware of a pain is one of several possible ways of knowing that one is in pain; and (iv) that it can be an open question, to be settled empirically, whether a person who is in pain and thinks he is in pain is entitled to say without

evidence that he is in pain.¹ Shoemaker, in fact, believes that the assertion 'I am in pain' does not require a justification. It is senseless, according to him, to ask the question 'How do you know that you have a pain?' because 'There is no logical possibility of my being unjustified in thinking that I am in pain.'² In this way he dispenses with introspection as a necessary condition of self-knowledge.

4.7 What are we to make of Shoemaker's argument for the rejection of the views that we observe or perceive our mental states, and that introspection is not a necessary condition of self-knowledge? His argument does not, I think, show that we do not, in any sense, observe or perceive our mental states. What his argument reveals is that if we do perceive or observe our mental states then such perception or observation cannot resemble our perception or observation of objects in the material world. In other words, his argument reveals that the traditional view of introspection as analogous to perception is mistaken, but not that there is no such thing as introspection. What Shoemaker fails to realise is that something like introspection does often seem to take place. We do, it seems, sometimes attend to 'the contents of our minds'. Consider, for example, such statements as 'My headache is getting worse' and 'My toothache is less excruciating now'. These statements suggest that we are, in some sense, paying attention to a headache and a toothache. Shoemaker does, however, succeed in establishing the important point that it is a mistake to think of ourselves as subjects perceiving or observing mental objects.

¹op.cit., p 223.

²op.cit., p 224.

Had he not succeeded in establishing this point the argument of the bundle theorists would have backfired immediately.

But if it is a mistake to think of ourselves as subjects perceiving our mental states, then is Hume not perhaps correct in thinking that a self is 'nothing but a bundle or collection of different perceptions'? I think not. Let me briefly explain why.

4.8 A serious difficulty with Hume's account of the nature of persons, in my opinion, is that there is no satisfactory explanation of how a variety of experiences are bound together in a single mind. Hume, it will be remembered, maintained that the experiences in a mind are related, in the same way that the discontinuous perceptions of a material object are, by resemblance, contiguity and causation. But to hold that experiences are related in this way is not to provide an adequate explanation of why we attribute a variety of experiences to the same self. For it is merely an empirical fact that experiences in a mind are related by resemblance, contiguity and causation; there is nothing to prevent there being a causal connection between my experiences and your experiences, e.g., it is not impossible that my experience of A always precedes your experience of B.

4.9 It is not surprising, therefore, that Hume ultimately rejected his account of why it is that we attribute a variety of experiences to the same self. In the Appendix to the Treatise he says 'but having thus loosened all of our particular perceptions, when I proceed to explain the principle of connexion, which binds them together, and makes us attribute to them a real simplicity and identity; I am sensible, that my

account is very defective.' Hume's dilemma was that, on the one hand, he did not see how the unity of the mind could be explained unless there was something simple and continued in which experiences inhere, and, on the other hand, he could not accept the existence of such an entity, on the grounds that all of our experiences are distinct and can be conceived as existing separately, and the mind never perceives any real connexion among distinct experiences.

4.10 It may be thought that Hume could have overcome his dilemma by holding that memory unites the various experiences of one's mind. But this does not seem to help. For memory seems to presuppose the existence of a persisting subject.¹ Firstly, it is not clear how a particular mental event could be remembered unless there was an enduring subject who first experienced the mental event and then later remembered having experienced the event. How else could a particular mental event and the remembering of the mental event, which is itself a mental event, be connected? Secondly, it is not clear how two memory experiences which occur at different times could be tied together, unless they were connected in virtue of the fact that they all belong to a persisting subject of experiences. The only way in which we can account for the unity of the mind, it seems, is by holding either that all our experiences belong to a pure ego, or that they are all causally dependent on the states of our body.

¹On this point see : A.J.Ayer, The Concept of a Person and Other Essays, Ch.4, pp.113-114.

4.11 Now even if the objection that it is impossible for Hume to offer a satisfactory explanation of the unity of the mind, could be met, his view would still be unacceptable. Exactly why it would still be unacceptable may be put as follows. If one holds that one consists of nothing more than a bundle of experiences then one must concede that when one goes to sleep, or is knocked unconscious, or is in a state of coma, one ceases to exist. But to say this is absurd; we do not regard going to sleep, or being knocked unconscious, or even being in a state of coma, as equivalent to death. This objection cannot, of course, be met by maintaining that one does not really cease to exist in such instances because one's experiences are causally dependent on the state of one's body which persists through sleep. For if it is true that one's experiences are causally dependent on the states of one's body then there is a possessor of experiences, viz., the body. Ayer, ⁽¹⁾ should be noted, maintains that 'a person's ownership of states of consciousness consists in their standing in a special causal relation to the body by which he is identified', but says that 'this amounts in effect to adopting what Mr Strawson calls 'the no-ownership doctrine of the self.'¹ He says this, it seems, because Strawson has, in providing a summary of the no-ownership doctrine, maintained that the idea of a possessor of experiences would have some validity, in the no-ownership

¹ op. cit., p 116.

theorists' view, if it was the body which was thought of as the possessor of experiences. That an out-and-out no-ownership theorist, like Hume or Wittgenstein, wouldn't agree with this, however, is obvious.

SUMMARY

5.1 Having shown that a self must be a subject of experiences we are in a position to consider, in the following chapters, whether a self is a material or immaterial subject of experiences. Wittgenstein, we have seen, rejected the view that the word 'I' refers to a person on the ground that there is no question of one's having to identify anything in order to say something like 'I am in pain'. I argued, in reply to this objection, that although the word 'I', like the word 'here', is, from the speaker's point of view, superfluous, the speaker uses the 'I' to refer to himself for the information of others. I pointed out, in support of this, that although one could say 'he is depressed' or 'So-and-so is depressed' when the person who is depressed is oneself, one says 'I am depressed' because one does not want to leave any doubt in the minds of others that it is oneself who is depressed. I then suggested that the 'I' in a sentence like 'I am in pain' could perhaps be said to indicate who is in pain, without presupposing the existence of a subject of experience, if persons were thought of as bundles of experiences; for then the 'I' could be said to indicate that the mind which is connected with the body from which the words 'I am in pain' are uttered is in a particular mental state.

I hope to have shown in the last few sections, however, that the Humean view of a person gives rise to insurmountable difficulties, and that it must, therefore, be discarded. I hope also to have shown that the bundle theorists' argument for the rejection of the view that a self is a subject is unsuccessful. The bundle theorists, it will be recalled, denied that experiences belong to a subject on the ground that nothing but mental states are observable in introspection. But what they failed to realise is that observation requires a subject as well as an object.

CHAPTER 2

STRAWSONIAN DUALISM

INTRODUCTION

1.1 One of the main difficulties with Strawson's account of what a person is, I suggested in the introductory chapter, is that it is not at all clear that he is entitled to hold that a person is a primitive type of entity. Bernard Williams in his review of Strawson's Individuals argues that Strawson's conception of a person is really Cartesian.¹ Strawson has, however, also been regarded by some as a behaviourist.² Strawson can, I think, easily seem to be a behaviourist in view of his claim that 'behaviour - criteria . . . are criteria of a logically adequate kind for the ascription of the P-predicate' (p 106). For it seems plausible to interpret this claim as implying that the relationship between behaviour and mental states is one of entailment. The trouble with this interpretation is that it is incompatible with Strawson's claim that one ascribes states of consciousness to oneself on a non-observational kind of basis. For how could one apply to oneself independently of observation a concept which is properly analysed exclusively in

¹ 'Mr Strawson on Individuals', Philosophy, 1961

² H.D.Lewis, for example, seems to regard Strawson as a behaviourist. See his 'Mind and Body - Some Observations On Mr Strawson's Views', P.A.S., 63 (1962-3), pp 1-22

terms of bodily behaviour? This leads us to another difficulty with Strawson's thesis. A.J. Ayer argues that it is impossible for Strawson to satisfactorily explain what he means by the expression 'logically adequate'.¹ It cannot, he maintains, imply that behavioural evidence entails the truth of the ascription of psychological predicates, for that does not stop short of physicalism. Nor can it imply that behavioural evidence provides sufficient empirical support for the truth of the ascription of psychological predicates, for that returns us to the argument from analogy, an argument which Strawson rejects in criticizing Cartesianism. But there is, Ayer concludes, no other way of interpreting Strawson's notion of a logically adequate criterion. His notion of a logically adequate criterion is thus incoherent. I shall discuss this criticism of Ayer's in the second half of this chapter. My aims in this chapter, then are (i) to determine whether Strawson's concept of a person really is primitive, and (ii) to find out whether it is possible for Strawson to offer a satisfactory explanation of the notion of a logically adequate criterion, i.e., an explanation which fits his conception of a person, whatever it is.

Is Strawson neutral between dualism and materialism ?

2.1 The ground for Strawson's distinction between M-predicates and P-predicates is, we have seen, that 'we would not dream of applying predicates ascribing states of consciousness' to material bodies. But what is meant by this? Bernard Williams in an article entitled 'Are Persons Bodies?' advances two possible interpretations of the claim that P-predicates are predicates we would not dream of applying to material bodies.² They are: (a)

¹ The Concept of a Person and Other Essays, Ch.4, pp 95-102.

² See Spicker ed., The Philosophy of the Body. Also reprinted in Williams' Problems of the Self.

that it would make no sense to apply such predicates to material bodies, and (b) that there are certain predicates which we ascribe to ourselves which we would never for a moment think could be truly ascribed to material bodies. Both interpretations are found by Williams to create serious difficulties for Strawson.

2.2 A more accurate paraphrase of interpretation (a), according to Williams, is that 'the conjunction of [the class of material body predicates] with any P-predicate does not make sense'.¹ The reason that this is a more accurate paraphrase, in his opinion, is because if anything is categorically senseless, conjunctions of predicates are . " 'That is green'", he says, "does not itself become a senseless form of words if someone tries to refer to a prime number in uttering the sentence . . . What surely are categorically senseless, if anything is categorically senseless, are sentences such as 'the prime number 7 is green'!"² But if this is right, Williams goes on to maintain, then it begins to look as though P-predicates and M-predicates cannot be jointly ascribed to anything, and hence that there are no persons. This conclusion would follow immediately if the class of material body predicates was 'coextensive with the class of M-predicates', or if 'any predicate ascribable to a material body can be sensibly, if patently falsely, ascribed to a person'. It is, however, impossible to characterise the class of material body predicates. According to Williams, 'there is a total obscurity about how these conjunctive sets of material-body predicates are in general to be characterised.' For it is impossible to say what 'in the required sense a material body' is. A material body cannot, he says, be merely anything to which M-predicates are ascribable,

¹ Problems of the Self, p.67.

² ibid.,

for then it would follow that there can be no persons. Nor can a material body be something to which just M-predicates are ascribable and P-predicates are not, for that 'gets us nowhere at all.' Having pointed out that there is no way of characterising the class of material body predicates, Williams advances an argument which, he thinks, does establish that if a P-predicate cannot be sensibly ascribed to a material body then the conclusion that there can be no persons does follow.¹ He concludes from all this that Strawson's thesis is hopeless if it is represented as a thesis about sense. If, however, interpretation (b) of the claim that P-predicates are predicates which we would not dream of applying to material bodies is the correct one, then, says Williams, a different objection will apply. The objection is that 'if it is just false of certain material bodies that psychological predicates apply to them, what obstacle can there be to saying that it is just true of others (e.g., ourselves) that such predicates do apply to them?'²

2.3 Williams' criticism seems to force Strawson to either acknowledge that his thesis is compatible with materialism, or, if he does not think it is, to say that it is not. If Strawson's view is that it is absurd to apply P-predicates to material bodies then, Williams has argued, it is difficult to see how P-predicates and M-predicates can be jointly ascribed to anything. But if Strawson's view is simply that we do not as a matter of fact ascribe P-predicates to material bodies then there is no obstacle to holding that persons are a class

¹ op.cit., p.69.

² op.cit., p.70.

of material bodies. So, assuming that Strawson's view is the latter view, he must either concede that his thesis is compatible with the view that persons are a class of material bodies, or defend the view that it is not. If he does think that there is an obstacle to holding that persons are a class of material bodies then this must, it seems, be because he thinks that a person consists of corporeal as well as incorporeal parts. But this conflicts with his claim that the concept of a person is primitive or unanalysable. So, if he is to be consistent, he must concede that his thesis is compatible with materialism. Strawson's claim that a person is an entity to which both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics are equally applicable is not, I have indicated, incompatible with materialism (for it is a topic neutral claim), though it would be if by 'states of consciousness' he meant non-physical states. But would Strawson concede that states of consciousness might be brain states? Would he concede that persons might be a class of material bodies?

2.4 What is most puzzling about Strawson's account of the nature of persons is that he seems to have three different referring expressions which could be substituted for the expression 'the very same thing' in the sentence 'One's states of consciousness are ascribed to the very same thing as certain corporeal characteristics'. These are 'a body' (p.89), 'a subject of experience' (p.93) and 'persons' (p.102). It is not at all clear, therefore, whether Strawson's account is coherent or not, or whether he is neutral between dualism and materialism.

Firstly, if the thing to which M-predicates and P-predicates are equally ascribable is 'a body', then it is surely contradictory to hold that 'we would not dream of applying predicates ascribing states of consciousness' to material bodies. Perhaps Strawson meant that we would not dream of ascribing states of consciousness to ordinary material bodies like rocks, trees, chairs, etc., and that states of consciousness can be truly ascribed only of complex material bodies like persons. But if this is what he meant (I very much doubt that he meant this), he should have said so. Secondly, the expression 'a subject of experience' is ambiguous, and can, therefore, be used either as an empirical concept, or as a non-empirical concept to denote a pure consciousness. If the expression 'a subject of experience' is used to denote a pure consciousness then, prima facie, it seems absurd to claim that M-predicates and P-predicates are ascribed to 'the very same thing', and that the concept of a person is primitive.

2.5 Now consider the following passage:

' . . . yet the facts in question still do not explain why we should, as we do, ascribe certain corporeal characteristics not simply to the body standing in this special relation to the thing to which we ascribe thoughts and feelings, etc., but to the thing itself to which we ascribe those thoughts and feelings' (p.93).

This passage suggests that a person is a possessor of a body. That Strawson does regard a person as a possessor of a body can also be seen in the first chapter in Individuals. He says,

right at the end of the chapter, 'It is a conceptual truth . . . that persons have material bodies' (p.58). Does this imply that Strawson is using the expression 'a subject of experience' to refer to a pure consciousness ? Not if we are going to take seriously his claim that a person is an entity to which M-predicates and P-predicates are equally ascribable. For if corporeal characteristics are ascribed to a person a person must be a material entity. But how can an entity which is the possessor of a body be itself a particular spatio-temporal entity, unless, of course, a person was identical with his body? According to Strawson, 'Persons, having corporeal characteristics, perceptibly occupying space and time, can be distinguished and identified, as other items having a material place in the spatio-temporal framework can be distinguished and identified' (p.132). But what room is there for the identification of a person as a different kind of entity to that of the material body which it possesses? How could we distinguish between a person as a possessor of a body and the body of a person? This interpretation of a person as a possessor of a body seems to me, therefore, to make hopeless Strawson's position. The only way in which Strawson can hold that 'a subject of experience' is a possessor of a body is by admitting either that a material body is part of a person, i.e., he has other non-corporeal parts as well, or that a person is identical with a material body. And this Strawson is not willing to do, for, he says 'The concept of a person is not to be analysed as that of an animated body or of an embodied anima' (p.103). We are also not, he says, to think

of a person 'as a secondary kind of entity in relation to two primary kinds, viz., a particular consciousness, and a particular human body' (p. 105).

2.6 There are some passages in Individuals which very much suggest, in my opinion, that Strawson's 'subject of experience' is 'a pure consciousness', and that his conception of a person is Cartesian. In his rejection of Cartesianism, for example, he says:

"One can ascribe states of consciousness to oneself only if one can ascribe them to others. One can ascribe them to others only if one can identify other subjects of experience. And one cannot identify others if one can identify them only as subjects of experience, possessors of states of consciousness" (p. 100).

It is clear that Strawson is using the terms 'subjects of experience' and 'states of consciousness' here to refer, respectively, to 'pure subjects' and non-physical states. So it seems that he regards a person as an entity consisting of 'a pure consciousness' and a body. Apart from the above passage, Strawson's talk of it being 'a contingent fact' that a subject of perceptual experience has just one body,¹ together with the admission that 'we might, in unusual circumstances, be prepared to speak of two persons alternately sharing a body, or of persons changing bodies' (p. 133), and such claims as 'the concept of a pure individual consciousness might . . . have a logically secondary existence' (p. 103), and 'each of us can quite intelligibly conceive of his or her survival of bodily death' (p. 115), also more than suggests that his conception of a person is Cartesian or near-Cartesian. With regard to the latter claim, it is I think true to say that only someone who thought that persons consisted of

¹ op. cit., p. 90.

material as well as immaterial parts, could 'intelligibly conceive of his or her survival of bodily death'.¹

2.7 So Strawson is not neutral between dualism and materialism. The reason that he claimed that the concept of a person is primitive is because it is 'logically prior' to the concept of an individual consciousness and the concept of a body. But this is not, it seems to me, a sufficient ground for holding that the concept of a person is unanalysable. How could Strawson come by the idea of 'a person's body' and 'the consciousness of that person' if the concept of a person were unanalysable? If the concept of a person was really primitive it would not be possible to hold that a person is a possessor of a body, and a particular of a different type, namely a consciousness. It seems to me, therefore, that what Strawson needs to do is to modify his claim that the concept of a person is 'logically primitive'. He should argue that although the concept of a person is initially unanalysable, there is nothing preventing us from analysing a person as an entity consisting of a mind (i.e., a non-physical entity) and a body once we have learnt how to identify a particular person. Strawson himself maintains that 'once we have identified a particular person, there is nothing to stop us, and nothing does stop us, from making identifying references to a particular of a different type, namely the consciousness of that person' (p 133). This is tantamount to holding that the concept of a person is analysable. Yet

¹ I am inclined to think that this is self-evident, but Bernard Williams has advanced an argument in support of this view in 'Are persons bodies', Problems of the Self, pp. 70-73.

Strawson, strangely, is not prepared to concede that a person may be analysed as an entity which consists of a mind and a body. If, however, he were prepared to admit that a person may be analysed in this way, he could hold that we ascribe M-predicates and P-predicates to the same thing only in a very weak sense. The difference between a Cartesian person and a Strawson person would be that whereas a Cartesian person is essentially a mind, a Strawson person consists of both a mind and a body, and is not to be identified with one or the other.

Can Strawson offer a satisfactory explanation of the notion of 'a logically adequate criterion'?

3.1 If Strawson is a dualist then it is not immediately clear what he means when he say 'one ascribes P-predicates to others on the strength of observation of their behaviour; and ... the behaviour-criteria one goes on are not just signs of the presence of what is meant by the P-predicates, but are criteria of a logically adequate kind for the ascription of the P-predicate' (p 106). He cannot mean that the behavioural evidence entails the truth of the ascription of psychological predicates, for that would put him in the position of the behaviourist. Nor can he mean that the behavioural evidence provides inductive support for the ascription of psychological predicates, for that can be so only because of an observed correlation between one's own mental states and one's behaviour. But what other possibility is there? Ayer expresses this difficulty as follows .

"The cardinal point is the attempt to stop short of physicalism on the one hand, and dispense with the argument from analogy on the other, by maintaining that our observations of the physical

condition and behaviour of other persons, on the basis of which we attribute experiences to them, are logically adequate for this purpose. But what exactly is meant here by saying that a criterion is logically adequate? Not that the evidence entails the conclusion, for in that case we should not stop short of physicalism . . . Not that the evidence provides sufficient empirical support for the conclusion, for then the reasoning is inductive; we are back with the argument from analogy. What is envisaged is something between the two, but what can this be? What other possibility remains?"¹

Ayer notes that Strawson is not the only philosopher to have held the view that there can be a relationship between statements which is in some sense logical but which is not deductive. Wittgenstein, he suggests, also held this view. But the sense in which it is held that the relationship between statements about behaviour and statements about experiences is logical is, for Ayer, a mystery.

3.2 It has been suggested to himself by Professor Alston, says Ayer, that the sense in which behavioural criteria are logically adequate for the ascription of experiences is that although they are not infallible, their overall success is logically guaranteed. The reason why their overall success is guaranteed is that without behavioural criteria we would be unable to talk significantly about our experiences. On this view, says Ayer :

"We are taught the use of a word like 'pain' in contexts in which the feeling for which it stands is outwardly manifested in some characteristic way; and the result is that this association is retained as part of the meaning which the word has for us." (p.100)

But even if this view is right it does not follow, in Ayer's

¹ The Concept of a Person and Other Essays, ch.4, p.95.

opinion, that we cannot dissociate these words, once our understanding of them has been acquired, from their original associations. He sees no reason why we cannot 'cut away the references to behaviour, and thereby obtain statements which were understood to refer to experiences alone.'

3.3 Before discussing Ayer's criticism it is worth noting that Strawson was not himself unaware that the statement 'behaviour criteria . . . are criteria of a logically adequate kind' etc. may seem to give rise to a difficulty. Almost immediately after having made this statement Strawson goes on to maintain that 'this is only one half of the picture about P-predicates.'¹ For one ascribes states of consciousness to oneself, according to him, independently of the behaviour criteria one uses in ascribing states of consciousness to others. He then asks how this fact is to be reconciled with the doctrine that behaviour criteria are logically adequate for the ascription of experiences to others. Exactly why there is a problem of reconciling the two doctrines he does not explicitly state. There is little doubt, however, as to why Strawson thinks we might have difficulty in holding both that behaviour criteria are logically adequate for the ascription of experiences, and that one ascribes experiences to oneself independently of observation of one's behaviour. If one holds that behaviour criteria are logically adequate for the ascription of experiences one denies the existence of 'a logical gap' between the criteria on the strength of which one ascribes a particular experience to another, and the actual experience. But this, it seems, is tantamount to accepting 'philosophical behaviourism', and is,

¹ op.cit., p.106.

therefore, incompatible with maintaining that one ascribes states of consciousness to oneself on a non-observational kind of basis. Strawson, however, does not think that there is a problem of reconciling two doctrines. According to him, it is essential to our understanding of these P-predicates that we recognize that there are two ways of ascribing them. It is essential to the single kind of meaning these predicates have that both aspects of their use, i.e., the self-ascriptive aspect and the other-ascriptive aspect, are acknowledged. It is a mistake, in Strawson's view, to think of any one aspect as primary or self-sufficient. If we take the self-ascriptive aspect of the use of these P-predicates as primary we open the way for 'philosophical scepticism'. And if we take the other-ascriptive aspect as primary we open the way for 'philosophical behaviourism'. But does Strawson's asserting that we must acknowledge both aspects of the use of P-predicates remove the difficulty involved in claiming that behaviour criteria are logically adequate for the ascription of experiences to others? It is not at all clear that it does. For if we are going to avoid 'philosophical scepticism' we must, it seems, interpret the claim that behaviour criteria are logically adequate etc. as implying that the relationship between behavioural criteria and experiences is one of entailment. But then we do not avoid 'philosophical behaviourism'. But perhaps there is another way of interpreting the claim that behaviour criteria are logically adequate etc. which will solve both this difficulty and the difficulty raised by Ayer.

3.4 Ayer's criticism of Strawson's notion of a logically adequate criterion appears, at first glance, to be insurmountable. It is my intention, however, to demonstrate

that this notion may be interpreted in such a way as to avoid Ayer's objection. Professor Alston's suggestion is, I believe on the right lines, though it clearly does require supplementation. The solution to Ayer's dilemma has, I shall argue, been provided, partly at least by Shoemaker's explanation of the notion of a criterion.¹ What has to be explained, for Ayer, it will be recalled, is the sense in which behaviour criteria are logically adequate for the ascription of experiences if the relationship between behavioural facts and mental phenomena is neither inductive, nor deductive. Shoemaker's explanation of the notion of a criterion enables us, in my opinion, to provide just such an explanation.

3.5 Shoemaker argues that 'we may characterize the criteria for the truth of a judgment as those states of affairs that are direct and non-inductive evidence in favour of the truth of a judgment.'² The assertion that a certain phenomenon is evidence in favour of the truth of a particular judgment is thus necessarily rather than contingently true. We know that a certain phenomenon is evidence for the truth of a judgment about X's identity, says Shoemaker, not by having noted correlations and discovered empirical generalizations, but by understanding the concept of X and the meanings of statements about the identity of X. Similarly, it could be said, we know that the display of pain-behaviour by Peter is evidence in favour of the truth of the judgment 'Peter is in pain', not by having noted correlations and discovered empirical generalizations, but by understanding the concept of pain.

¹ In his Self-knowledge and Self-Identity, pp.3-4.

² op. cit., p.3.

Peter's pain-behaviour may be said, therefore, to constitute 'direct and non-inductive evidence' in favour of the truth of the judgment 'Peter is in pain'. But how, it may be asked, does our possessing the concept of pain entitle us to say that it is a necessary truth that Peter's pain-behaviour is evidence for the truth of the judgment 'Peter is in pain'? By way of answering this question I want now to turn to an article of Robert Coburn's entitled 'Persons and Psychological Concepts.'¹

3.6 Coburn also maintains that we must understand the notion of a criterion in the way in which Shoemaker suggests. We are to understand by the notion of 'a logically adequate criteria' for the application of a psychological concept, he says, 'a phenomenon or circumstance (or set of phenomena or circumstances) which constitutes evidence that the concept in question applies in virtue of the very structure of the concept, and which hence constitutes "noninductive evidence" that the concept in question applies.'² He subsequently expands on, and clarifies the above by calling 'to mind the general point made by Wittgenstein that the language games we play (or the concepts we employ) "rest upon" certain contingent facts concerning human beings and/or the world in the sense that if certain facts were otherwise, it would be logically impossible for us to play these language games (or employ these concepts).'³ If the existence of a certain regular connection between behavioural facts and a mental phenomenon like being in pain were a logically necessary

¹ American Philosophical Quarterly, Volume 4, 1967, pp.208-221.

² op. cit., pp.210-211.

³ op.cit., p.212.

condition of our possession our concept of pain, then clearly, Coburn maintains, the behavioural facts would constitute non-inductive evidence for the application of the concept of pain. He proceeds to demonstrate that the evidence of certain general correlations between behavioural facts and mental phenomena is a logically necessary condition of our possessing psychological concepts, by advancing an argument which is 'suggested in part by certain remarks of Strawson'¹ If we never know on the basis of observation that psychological concepts are applicable to others, then it would seem to follow, he says, that we never know in any way that such concepts are applicable to others. We could have no idea of the occasions on which we could justifiably claim to know that such concepts apply to others. But if this were true 'it would seem that we could not even apply such concepts to ourselves. This is because, as Strawson puts it, "The idea of a predicate (a concept) is correlative with that of a range of distinguishable individual of which the predicate (or concept) can be significantly, though not necessarily truly, affirmed(applied)" (p.99,n.i).'²

Unfortunately, Coburn does not attempt to explain why the idea of a predicate (or concept) should involve the idea of a range of individuals to which the predicate (or concept) can be applied. Strawson's claim that the idea of a predicate involves the idea of a range of individuals to which the predicate can be significantly affirmed seems to rest, however, on the belief that it is, in

¹ op.cit., p. 215.

² ibid.

principle, impossible to acquire the possession of a private language. The reason why Strawson thinks that it is impossible to have a private language is, it seems, because he thinks that it would be impossible to have the idea of a possessor of experiences unless we had learnt how to ascribe experiences to others. A necessary condition of ascribing psychological predicates to oneself, in Strawson's view, is that one has the idea of some particular thing to which these predicates are ascribable, i.e., of a possessor of experiences. But one can have no such idea, he says, unless one has learnt how to ascribe experiences to others. The unique position of a single body in one's experience does not in itself, in Strawson's view, explain why experiences should be ascribed to something which has them. And we can have no 'inner intuition' of an immaterial subject of experiences, as both Hume and Kant have observed. We cannot, therefore, say such things as 'I have a toothache', 'I see a spider on the ceiling', 'I am in pain', etc., unless we have learnt to ascribe experiences to others, and so acquired the idea of a possessor of experiences.¹

3.7 Coburn substantiates the above argument by mentioning a number of absurdities which follow from the claim that such concepts as the concept of pain can never be justifiably other-applied.² Three of the absurdities he mentions are the following : (i) 'we could never be sure that such concepts are inapplicable to any things other than ourselves';

¹ I shall expand on these remarks in the chapter on Cartesian dualism.

² op. cit., p.217.

(ii) 'we could never be sure that anyone else understands the meanings of words like "pain", "belief", etc.'; (iii) 'no one could teach another the meanings of words like "pain", "believe", and the like.' In view of these absurdities, and the argument in the previous paragraph, it is, I think, clear that the presence of behavioural facts do constitute 'noninductive evidence' for the application of psychological concepts.

3.8 So we can meet Ayer's objection by maintaining that the sense in which behaviour criteria are logically adequate for the ascription of experiences is in that they constitute direct and noninductive evidence for the ascription of experiences. Contrary to what Ayer believes, the connexion between behaviour criteria and experiences is a necessary one even though a certain type of behaviour does not entail a particular experience. Behavioural facts constitute noninductive evidence for the ascription of experiences, we have seen, because it is a logically necessary condition of our having psychological concepts that there is a regular connection between behavioural facts and mental phenomena. Our understanding of the concept of pain, for example, necessitates that we grasp the connexion between pain and pain-behaviour. One may, of course, be mistaken in one's ascription of a particular experience to another. Strawson himself admits that 'what can be observed can also be faked or disguised'.¹

¹ op .cit., p.109.

But this clearly does not open the way for 'philosophical scepticism.' For in the absence of behaviour criteria it would be impossible to ascribe experiences at all. As Wittgenstein put it 'An inner process stands in need of outward criteria.'¹ Furthermore, the view that all relationships between psychological states and physical states can be contingent is obviously incoherent. If this were the case our psychological concepts could not have established meanings, for it would be impossible for people to be taught their meanings.² Although pain-behaviour is not, therefore, a necessary or a sufficient condition of pain it is necessarily good evidence for pain.

3.9 Now although I have defended Strawson against Ayer's objection that he can give no satisfactory explanation of the claim that behaviour criteria are criteria of a logically adequate kind for the ascription of psychological predicates, it is clear that it would, indeed, be impossible to make sense of this claim if we interpreted 'logically adequate', as Ayer thinks we should, to mean deductively adequate. If we interpreted the expression 'logically adequate' in this way then behaviour criteria could be criteria of a logically adequate kind for the ascription of psychological predicates only if the relationship between behaviour and mental phenomena is one of entailment. The behaviour criteria would, in other words, have to guarantee that a person behaving in a certain way was

¹ Philosophical Investigations, Section 580.

² This point is made by Shoemaker, op.cit., p.169.

in a particular psychological state, just as the premises of a valid deductive argument must guarantee or entail the conclusion of the argument. Now strictly speaking it may be correct to interpret 'logically adequate' in this way, but this is not the only way of interpreting this expression. What Strawson means when he says that behaviour criteria are criteria of a logically adequate kind for the ascription of psychological predicates is that behaviour criteria provide necessarily good evidence for the ascription of psychological predicates, and they provide necessarily good evidence in virtue of the very structure of the concepts, for it is a necessary condition of our understanding psychological concepts that we grasp the connexion between mental phenomena and behaviour. For Strawson, then, the expression 'logically adequate' means conceptually adequate. Whether Strawson is correct or not, therefore, in using the expression 'logically adequate', it is not the case, as Ayer thinks, that it is impossible for Strawson to offer a satisfactory explanation of what he means by the claim that behaviour criteria are adequate for the ascription of mental states. It is not the case that it is impossible to make sense of this claim unless this amounts to the claim that behavioural evidence entails the truth of the ascription of psychological predicates, or the claim that the behavioural evidence provides inductive support for the truth of the ascription of psychological predicates. For, to repeat, the sense in which behaviour criteria are criteria of an adequate kind for the ascription of experiences is in the sense that they constitute 'direct and noninductive evidence' for such ascription.

SUMMARY

4.1 Strawson's view of a person is, I have argued, very similar to Descartes' in that he also seems to think of a person as an entity consisting of two distinct parts, viz., a body and a non-physical mind. His claim that a person is an entity to which both M-predicates and P-predicates are equally applicable does not in itself, we have seen, suggest that his view of a person is similar to the Cartesian dualists, for, as Williams' argument has brought out, it is compatible with materialism to hold that two types of predicate are ascribable to persons. That Strawson's view is very similar to Descartes' is, I have argued, revealed by his use of the expression 'a subject of experience' and by the presence of such claims as 'each of us can quite intelligibly conceive of his or her survival of bodily death'. His claim that behaviour criteria are logically adequate for the ascription of psychological predicates does not, we have seen, conflict with his essentially Cartesian view of a person, since the claim that behaviour criteria are logically adequate etc. does not necessarily imply that the relationship between a particular piece of behaviour and a particular mental state is one of entailment. But the relationship between a particular piece of behaviour and a particular mental state is, I have argued, a necessary one, in virtue of the fact that behaviour criteria constitute 'direct and noninductive evidence' for the ascription of psychological predicates. Behaviour criteria constitute

'direct and noninductive evidence', we have seen, because it is a necessary condition of our possessing our psychological concepts that we grasp the connexion between mental phenomena and behavioural facts. But if possessing psychological concepts involves grasping the connexion between mental phenomena and behavioural facts it would seem to follow that part of what we must mean when we say something like 'So-and-so is in pain' is that he is behaving in one of a certain number of ways.

CHAPTER 3

CARTESIAN DUALISM

INTRODUCTION

1.1 Strawson's argument for the rejection of the Cartesian view of a person is to be found in Section 4. of the chapter 'Persons' in Individuals. Not all philosophers, I have maintained, have been convinced by his argument. H.D.Lewis attempts to counter Strawson's attack on Cartesianism in a chapter entitled 'Persons and The Structure of Language' in his book The Elusive Mind.¹ Lewis' arguments in this chapter are exceedingly bad, largely because he has failed to grasp Strawson's position properly, but they do help to illuminate Strawson's own argument. Lewis advances three main objections. He begins by questioning whether it is strictly the case that we ascribe both corporeal characteristics and mental characteristics to the same thing (pp 148-151). He then proceeds to dispute the view that 'we ascribe states of consciousness to ourselves in precisely the same way as we ascribe them to others' (pp 153-155). Finally, he attacks the view that it would be impossible to ascribe states of consciousness to others if they are thought of as a set of Cartesian egos (pp 155-159). It is my intention to show that each of these objections is either misguided or mistaken. Having done this I shall then attempt to show that

¹This chapter also appears as an article entitled 'Mind and Body - Some Observations on Mr Strawson's Views', in Proceedings of the Aristotelian Society (1962-3) pp.1-22.

Cartesianism is not entirely defenceless against Strawson's attack. In attempting to defend Cartesianism I shall refer to Locke's view of a person. Finally, I shall argue that even though Strawson's argument fails there are still good reasons for rejecting the Cartesian view of a person. But we must first see why Strawson rejected this view of a person.

Strawson's objection to Cartesianism

2.1 Strawson's chief objection to Cartesianism may be put as follows:¹ A necessary condition of ascribing experiences to oneself is that one should also ascribe them, or be prepared to ascribe them, to others; a necessary condition of ascribing experiences to others is that one can identify other subjects of experience; other subjects of experience could not be identified if they were only subjects of experience, i.e., possessors of consciousness; a necessary condition of identifying other subjects of experience, therefore, is that they are entities to which both states of consciousness and corporeal characteristics are equally ascribable.

2.2 The problematical premise is the first one. It suggest that in order to ascribe experiences to oneself, there need not actually exist other subjects of experience to which one can ascribe experiences, but that one must at least recognize the logical possibility of ascribing experiences to other subjects.² In order to have an idea of myself as a subject of experience I must at least be able to imagine other subjects of experience with which I can contrast myself. Strawson

¹ See Individuals, pp.99-100. Also see p.104.

² This is suggested, in particular, by the expression 'or be prepared to ascribe them.'

indicates a couple of pages further on, however, that he does not think we could imagine ascribing experiences to others unless we could actually identify other subjects of experience. He says, 'There is no sense in the idea of ascribing states of consciousness to oneself, or at all, unless the ascriber knows how to ascribe at least some states of consciousness to others'¹. It might be possible to object here (owing I think to the ambiguity of the claim) that to hold that it does not make sense to ascribe experiences to oneself unless there are other identifiable subjects of experience is not the same as holding that it is impossible to conceive of oneself as a subject of experience unless there are other identifiable subjects of experience. The reason why, it might be held, it does not make sense to ascribe experiences to oneself is simply because there is only one subject of experience. There is no need for a subject who is unaware of the existence of other subjects to refer to his experiences as 'mine' or to say such things as 'I am in pain', 'I feel depressed', etc. It is plain, however, from other things which he has to say, that Strawson does think that in order to have an idea of oneself as a subject of experience one must actually be able to identify other subjects of experience. He says, for example, that 'states of consciousness could not be ascribed at all, unless they were ascribed to persons in the sense' which he has claimed for the word.²

2.3 On what grounds does Strawson hold that it is impossible to ascribe experiences to oneself unless one already

¹ op.cit., p.106.

² op. cit., p.102.

knows how to ascribe them to others? The answer to this question may be put as follows: The claim that it is a necessary condition of ascribing experiences to oneself that one should also ascribe them to others rests on the assumption that it is impossible for a person to acquire a private language for himself. Strawson does not deny that it is possible for a person, who is already a common language user, to invent a private language for himself - he admits that 'the idea of a uniquely applicable predicate, i.e., a predicate which belongs to only one individual, is not absurd' (footnote, p 99) - but he does deny that it is possible for a person who is not a common language user to ascribe predicates to himself. The reasoning behind this denial is that whereas the common language user, in the process of being taught how to ascribe predicates to others, has acquired the idea of a possessor of experiences, the individual who is not a common language user can have no idea of any particular thing to which experiences can be ascribed. He can have no idea of any particular thing as the 'possessor' or 'owner' of experiences. Strawson, it should be noted, would not deny that such an individual could have concepts. He would concede that such an individual could have a concept of, say, pain, but he would deny that he could have a concept of pain which is like our concept of pain, that is, which involves the idea of a possessor of pain. He says, in his review of Wittgenstein's Philosophical Investigations, 'Wittgenstein gives himself considerable trouble over the question of how a man would introduce a name for a sensation into this private



language. But we need imagine no special ceremony. He might simply be struck by the recurrence of a certain sensation and get into the habit of making a certain mark in a different place every time it occurred.¹ Such an individual could have a concept of a sensation like pain, though his concept of pain would not involve the idea of a possessor of pain. A person who has not been taught a language could, in Strawson's view, have only a concept of what could be called unasccribed pain.

2.4 Why, it might now be asked, can an individual, who is not a common language user, not imagine what ascribing experiences to others would involve, and so acquire the idea of himself as a possessor of experience? Why is it that he can have no idea of a particular thing to which experiences are ascribable? Strawson's answer to these questions is perhaps not all ^{that} clear, yet it is plain that he does have an answer. Firstly, the fact that 'for each person there is one body which occupies a certain causal position in relation to that person's perceptual experience' does not in itself, in Strawson's view, explain why thoughts and experiences should be ascribed to any subject.² And, secondly, we can have no empirical awareness of an immaterial object which is the unitary subject of experiences. This second point is not made explicitly, but is, I think, implicit in what Strawson says. It derives from Kant, and it is therefore surprising that Strawson does not refer to Kant's refutation of Cartesianism.³

2.5 Let us briefly turn our attention to Kant's argument. Kant, unlike Strawson, does not argue that one can

¹ Mind, Vol. 63, 1954, p. 85.

² op.cit., pp. 92-93.

³ Kant's refutation of Cartesianism is to be found in the section of The Critique of Pure Reason entitled 'Paralogisms of Pure Reason'.

have no idea of oneself as a subject of experience unless one is able to identify other subjects of experience. He argues simply that we cannot infer from the proposition 'I think' that we exist as a Cartesian thinking substance. 'No mere analysis of the proposition "I think"', says Kant, 'will suffice to prove such a proposition.'¹ The 'I' of which one can be conscious in every act of thought does not, he points out, involve any intuition (empirical awareness) of a thinking substance. Consciousness of oneself in thought, he says, 'does not concern any intuition of the subject, whereby it is given as object, and cannot therefore signify the identity of the person, if by that is understood the consciousness of the identity of one's own substance, as a thinking being.'² In order to claim knowledge of any such subject it must, Kant maintains, be known by direct acquaintance. For if we can have no intuition of an immaterial thinking substance 'we cannot know whether there is any object to which the concept is applicable . . . (and) the concept therefore yields no knowledge whatsoever.'³ Kant concludes from the fact that we can have no intuition of the 'I' of which we are conscious in thinking that it is 'an entirely empty expression', that it is a 'mere logical subject' of experiences. He acknowledges that it is through the 'I' that we have knowledge of the unity of our experiences - all of one's experiences can be ascribed to a logically simple subject - but claims that it does not therefore follow that we have knowledge of an immaterial thinking substance. He says, 'The identity of the consciousness of myself at different times is therefore only a formal condition

¹ B 408.

² B 408.

³ B 412.

of my thoughts and their coherence, and in no way proves the numerical identity of my subject.¹

2.6 Strawson, it is worth noting, argues in The Bounds of Sense, which is an examination of Kant's Critique, that it is implicit in Kant's position that ascribing experiences to a subject requires empirically applicable criteria of identity, and that one cannot ascribe experiences to oneself, therefore, unless one possesses the concept of a man or human being. One of the weaknesses in Kant's exposition, says Strawson, is that his reference to the fact that our ordinary concept of a person carries with it empirically applicable criteria of subject - identity, is minimal and obscure. Kant alludes to this point, according to him, in the sentence: 'Its (the soul's) permanence during life is of course evident, since the thinking being (as man) is itself likewise an object of outer senses.' This point, Strawson stresses, is 'of the first importance' for 'It means that we have, after all, a concept, which satisfies the most stringent critical requirements, of a persisting subject of experiences (a man). This concept supplies an absolutely firm basis for a genuinely object-referring use of personal names, and of persons and pronouns, in sentences in which states of consciousness, inner experiences, are ascribed to the objects referred to by names or pronouns.'² Without such a concept it would, Strawson has maintained, be impossible to ascribe experiences at all.

¹ A363; see also B 409.

² The Bounds of Sense, p. 164.

2.7 Let me now sum up what I have said concerning Strawson's claim that it is a necessary condition of ascribing experiences to oneself that one should be prepared to ascribe them to others. The reasoning behind this claim is, I have suggested, as follows : A condition of ascribing experiences to oneself is that one has a grasp of the contrast between oneself and others as subjects of experience. A condition of one's grasping the contrast between oneself and others is that one can identify other subjects of experience. One cannot imagine ascribing experiences to others because one can have no idea of a possessor of experiences at all unless we can actually identify other subjects of experience. The possession of a body to which one has a special attachment does not in itself provide any good reason for ascribing experiences to a subject. And we can have no intuition of an immaterial object which is the unitary subject of experiences. One can have no idea of any particular thing to which experiences can be ascribed, therefore, unless one possesses the empirical concept of a subject of experience, viz., the concept of a man or human being.

Lewis' defence of Cartesianism

3.1 Having attempted to elucidate Strawson's argument for the rejection of Cartesianism we may now turn to consider Lewis' defence of Cartesian dualism. Lewis begins his defence of Cartesian dualism by arguing that Strawson's claim that we ascribe states of consciousness and physical

characteristics 'to the very same thing' is false. He admits that we ascribe states of consciousness and physical characteristics to the same thing 'in a rough and ready way for ordinary purposes', but maintains that this is because "It would be much too troublesome to say, for example - 'I was intending to open the door and my body moved towards it'. It is always neater, and apter, for ordinary purposes, to say simply 'I went to the door' or 'I went to open the door'".¹ Strictly speaking, says Lewis, we do not ascribe physical characteristics and states of consciousness to the same thing. In his view we ascribe physical characteristics to the body and states of consciousness to the mind. He says "Where, for example, I say 'I am tall', I am not saying anything about my mind but only about my body".² Again, further on, he says 'The strict truth is not that I am bald, although that is a perfectly clear way of putting it for normal purposes, but that my head is bald. It is my own head, part of a body to which I stand in a very special relation, but my mind is neither bald nor covered with hair.'³ Lewis, it can be seen, identifies himself essentially with his mind.⁴ What are we to make of his argument against Strawson? He is, I think, right to insist that it does not follow from the fact that we say such things as 'I have black hair and green eyes' and 'I have a toothache' that we are referring to the same thing in both of these sentences. For neither the claim that the word 'I' is univocal nor the claim that the word 'I' is equivocal appear to be obviously true or false. What Lewis fails to realise, however, is that Strawson's position is really very similar to the Cartesian theorists'

¹ op.cit., p.148.

² ibid.

³ op.cit., p.151.

⁴ He actually says in the next paragraph that 'my real self is my mind'.

position (his own position). Strawson is, I have argued, ultimately committed to the view that a person is an entity consisting of logically distinct parts, viz., a body to which M-predicates are applicable and a consciousness to which P-predicates are applicable. He is, in other words ultimately forced to reject the view that a person is a logically primitive bearer of both P- and M-predicates. Strawson would, therefore, clearly agree with Lewis' claims that when he says 'I am tall' or 'I am bald' he is not saying anything about his mind, but only about his body, though he would deny that he is not saying something about a person. Descartes himself, it is interesting to note, also sometimes uses the word 'I' to refer to an entity consisting of both a mind and a body, though he does also sometimes use the word 'I' to refer simply to the mind.¹ Lewis, it seems, thinks that Strawson's claim that we ascribe physical characteristics and states of consciousness to the same thing implies that the thing to which these characteristics apply is a mere material thing. This, however, is clearly a mistake.

¹ In his Second Meditation Descartes asserts 'I now admit nothing that is not necessarily true: I am, therefore, precisely speaking, only a thinking thing, that is a mind'. In his Sixth Meditation, however, he maintains 'Nature likewise teaches me by these sensations of pain, hunger, thirst, etc., that I am not only lodged in my body as a pilot in a vessel, but that I am besides so intimately conjoined, and as it were intermixed with it, that my mind and body compose a certain unity.

3.2 Lewis' next object, it will be recalled, is that Strawson is mistaken in thinking that we 'ascribe states of consciousness to ourselves in precisely the same way as we ascribe them to others'.¹ The expression 'in the same way', it should be noted, is ambiguous in that it can suggest either that the ascribing phrases have the same meaning whether one is ascribing experiences to oneself or others, or that the method of ascribing experiences to oneself and others is similar. Lewis, however, is using the expression in the latter sense. He attributes Strawson with the view that the ascribing phrases could not have the same meaning unless the method of ascribing experiences to oneself and others is similar.² Hence, according to Lewis, Strawson's view that we can 'only identify ourselves when we also identify others' (p.154). This, he says, together with the fact that we can only identify persons through observing their bodies, suggests that it is 'impossible to ascribe experiences to oneself at all except in ways in which our bodies have an indispensable part.' (p 157). It should be apparent by now that this view definitely cannot be attributed to Strawson, and that Lewis' second objection is therefore entirely misguided. Strawson's view, as we have already seen, is that one ascribes states of consciousness to oneself independently of the behaviour criteria one uses in ascribing states of consciousness to others. He also maintains that one ascribes predicates such as 'going for a walk', 'coiling a rope', 'playing ball', etc., to oneself independently of observation of one's behaviour, and that such predicates

1 op.cit., p.154.

2 ibid

release us from the idea that the only things we can know about without observation or inference, or both, are private experiences.¹ In what way then does ascribing experiences to oneself involve the identification of a body?

3.3 Lewis attributes to Strawson the view that we can 'only identify ourselves when we also identify others', it seems, because of the way in which he interprets Strawson's claim that 'One can ascribe states of consciousness to oneself only if one can ascribe them to others'. He appears to take this to mean that one can have an awareness of one's experiences only in a way that directly involves the awareness we have of the experiences of others. The point of Strawson's claim, we have seen, is that it would be impossible to ascribe experiences to oneself unless one already knew how to ascribe them to others. It is impossible, in Strawson's view to arrive at a conception of oneself as a subject of experience merely through the awareness we have of our experiences. Lewis, as we shall see from his rejection of Strawson's reply to a possible defence of Cartesianism, thinks that Strawson is mistaken in this. He says, rather simplistically, 'I know that a private experience is mine in having it'. The important point to note for the moment, however, is that Strawson's saying that one cannot ascribe experiences to oneself unless one has been taught how to ascribe them to others, is quite different from saying that one must always have an

¹ op.cit., p.111.

awareness of others when one is self-ascribing experiences, or that one ascribes experiences to oneself in precisely the same way (i.e., using the same method) that one ascribes them to others.

3.4 One of the things which seems to have lead to Lewis' misunderstanding of Strawson's position is the account which Strawson gives of the concept of depression. We speak of depressed behaviour and we also speak of a feeling of depression, but there is, says Strawson, no room here to drive in a logical wedge. The reason why there is no room to drive in a logical wedge is because, in Strawson's own words, 'X's depression is something, one and the same thing, which is felt, but not observed, by X, and observed, but not felt, by others than X' (p.109). Lewis claims that he finds this argument very mystifying. I cannot see why he should. There is, as we saw in the chapter on Strawsonian Dualism, no logical gap between behaviour criteria and experiences because it is a necessary condition of ascribing experiences that there are behaviour criteria. Our understanding of our psychological concepts necessitates that we grasp the connection between experiences and behaviour. It is thus part of the meaning of 'depression', that an individual who exhibits a certain type of behaviour, viz., depressed behaviour, is said to be in a state of depression. The ascribing phrase 'is depressed' means the same thing, therefore, whether one says 'I am depressed' or 'He is depressed'. It is for this reason, I take

it, that Strawson claims that 'X's depression is something, one and the same thing,' etc.

3.5 Lewis' final objection deals specifically with Strawson's criticism of the Cartesian position. He attacks the view that there is no way of telling that a private experience is another's if they are thought of as a set of Cartesian egos. He begins by maintaining, mistakenly, that this view rests on the unwarranted assumption that if experiences are private to those who have them there is no way in which they can be made public or known indirectly. Now Ryle may make this assumption, but Strawson certainly does not. Strawson's view, in fact, is that we do not have direct access to the states of consciousness of others, and that we do infer their state of consciousness from their behaviour. He never says as much, however, for he does not think we could arrive at this conclusion unless we have learnt how to ascribe states of consciousness to others. It is for this reason that he argues that behaviour is more than a 'sign of the presence' of a state of consciousness. Only after we have learnt how to ascribe states of consciousness to others is it possible to note that we have a privileged access to our own experiences.

3.6 The fact that Lewis rejects Strawson's reply to a possible defence of Cartesianism indicates, in my opinion, most conclusively, that he has not properly grasped Strawson's initial objection to Cartesianism. The possible defence of Cartesianism, which Strawson considers and then rejects, is that there is no difficulty in distinguishing

bodies from one another and that this gives us an indirect way of identifying subjects of experience. He rejects this defence on the grounds that 'It requires me to have noted that my experiences stand in a special relation to body M, when it is just the right to speak of my experiences at all, that is in question' (p 101). Lewis is not persuaded by this argument. According to him, "The 'my' gets into the real argument on the basis of one's experience of oneself as a conscious being."¹ It is clear, however, that such a response will not do. For simply to assert that one is entitled to employ the possessive pronoun 'my' on the basis of one's experience of oneself as a conscious being is to beg the very question at issue, viz., that one can have no conception of oneself as a subject of experience unless one knows how to identify other subjects of experience. Strawson would agree with Lewis that it is a fundamental requirement of the possibility of empirical self-consciousness that we have an awareness of our various experiences, but, he would point out, this is not a sufficient requirement. A further requirement for the actual ascription of experiences to oneself is that one possesses 'the empirical concept of a subject of experience'. In order to have a concept of a subject one must have empirical criteria of subject-identity, but none such are supplied, to use Kant's terminology, in 'inner sense'. We can, for example, have no 'inner intuition' of a persisting immaterial subject of experiences. If I cannot identify other subjects of experience, therefore, I cannot argue from 'my own

¹ op.cit., p.158.

case' that the identifying of bodies is an indirect way of identifying subjects of experience. I could have no conception of myself as a conscious being and, therefore, could not note 'that my experiences stand in a special relation to body M'. This is not to say, however, that I do not have what could be called a 'native awareness' of my experiences. Animals and infants presumably have just such an awareness of their experiences.

A reply to Strawson's objection

4.1 So much for Lewis' defence of Cartesianism. Although Lewis' defence fails, the Cartesian view is not, in my opinion, entirely defenceless against Strawson's attack. If the claim that 'One can ascribe states of consciousness to oneself only if one can ascribe them to others' is true, and I have indicated that I think it is, then, clearly, it creates a strong presumption in favour of the view that persons are essentially corporeal beings. But it does not prove that persons are corporeal beings. The Cartesian theorist can agree with Strawson on the way in which we acquire the conception of ourselves as subjects of experience, but still insist that the view that a person is essentially a corporeal being is mistaken. For the most part the facts that it is a necessary condition of ascribing experiences to oneself that one can ascribe them to others, and that it is a necessary condition of ascribing experiences to others that they are entities to which corporeal characteristics are also ascribable, could be said to establish, is that persons must either be material entities or they must be embodied. These facts do not, therefore, conclusively prove

that persons must be material entities; they leave open the possibility that persons could be identified indirectly by way of their bodies. And Strawson himself does not maintain that persons are only corporeal beings. Persons, in his view, we have seen, consist of bodies as well as non-physical minds. If the Cartesian view is that a person is an entity consisting of both a body and a mind, therefore, there is little difference between the Strawsonian and Cartesian conceptions of a person. But even if the Cartesian view is that a person is to be identified solely with a mind, it does not follow that it cannot be defended from Strawson's attack. The Cartesian theorist can admit that he could not have a concept of himself if persons did not have bodies, but still reject the view that a person is essentially a corporeal being. For if one accepts, say, Locke's definition of a person as 'a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking, in different times and places', it is not immediately apparent that the Cartesian theorist is mistaken in claiming that persons are essentially immaterial entities.¹

4.2 Locke's definition of a person, it can be seen, leaves open the question of whether persons are material or immaterial entities, though Locke himself thought persons were essentially noncorporeal entities. He distinguished the

¹ See John Locke, Essay Concerning Human Understanding, Book 2, Ch.27, Section 9.

idea of a man from the idea of a person. The idea of a man, according to him, involves the idea of 'a body so and so shaped', while the idea of a person involves the idea of 'a thinking or rational being'. A rational parrot would not, therefore, be called a man, but it could be called a person. Any animal endowed with reason and reflection could, on Locke's account, be called a person. The defining characteristic of a person is not that he is shaped in a certain way, but that he is capable of reason and reflection. This characterization of a person is, in my view, a highly plausible one, for we do tend to think that the fundamental difference between persons and other material bodies is that whereas the former are capable of reason and reflection, the latter are not.¹ But if one accepts Locke's characterization it is not, I have suggested, immediately clear that it is absurd to hold that persons are noncorporeal entities.

4.3 Strawson may, however, want to defend his position, for he does argue that a subject of experience existing apart from a body could think of himself only as a former person. A subject of experience existing apart from a body must, he says ,

"contrive still to have the idea of himself as a member of a class or type of entities with whom, however, he is now debarred from entering into any of those transactions the past fact of which was the condition of his having any idea of himself at all" (p.116).

¹ Locke's characterization of a person actually has a distinct advantage over Strawson's, in that it does not restrict the class of persons to the class of human beings. On Strawson's view, if a particular entity is incapable of smiling or going for a walk that entity is disqualified from being regarded as a person. The problem with restricting the class of persons to the class of human beings, however, is that it may well happen that certain animals, like dolphins, will in time come to be regarded as persons. On this point see: Moulder, J., 'In Defence Of Immaterial Persons' *Philosophical Papers* May 1972 pp50-52

But why this should be so is not at all obvious, as I shall shortly reveal. Strawson also maintains that a disembodied person will, as his memory fades and his interest in human affairs dwindles, cease to have an idea of himself as an individual. 'In proportion as the memories fade, and this vicarious living palls, to that degree his concept of himself as an individual becomes attenuated'. But why this should be so is also far from clear. Let us take this latter point first. A disembodied person, on Strawson's account of survival, can observe other persons, have thoughts about other persons, even have feelings towards other persons. Why, then, should his memory fade when, in view of the fact that he can observe people acting in precisely the same way he once acted, he has a constant reminder of himself as a former person? And surely there is no chance of his interest in human affairs dwindling, for as a former person he will have an intense interest in the destiny of the human race.

4.4 Strawson's account of survival also makes it difficult to see why a disembodied individual must always think of himself as a former person. A disembodied person, Strawson suggests, need not live solely in a world of memories. It is more than likely, therefore, that most disembodied individuals would, in view of their detachment from the material world, alter their opinions on certain matters, adopt different attitudes to various persons they had known, acquire new views on a variety of subjects, and thus continue to broaden and deepen

their experience. Some disembodied individuals might even realise some latent capacity, such as the capacity to appreciate works of art, or the capacity to appreciate classical music. These individuals might well, thereafter, spend a great deal of their time frequenting art galleries and concert halls, as unknown art enthusiasts and lovers of classical music. Besides these considerations, there is the further consideration that disembodied individuals might be able to communicate with one another. The possibility of communications between disembodied individuals is not, it seems, ruled out on Strawson's account of survival. Strawson does, admittedly, exclude the possibility of a disembodied person communicating with an embodied person, though he does not advance any reason for excluding this possibility. If, however, the notion of telepathic communications between embodied persons is not an incoherent one, then I do not see why, on Strawson's account of survival at any rate, the notion of telepathic communication between disembodied persons should be. Strawson should have little difficulty in imagining two disembodied individuals tapping each other's minds.

4.5 The claim that a disembodied person could think of himself only as a former person is not, therefore, well-founded. The reason that Strawson introduces this claim, I have suggested, is that he thinks it enables him to hold, without fear of contradiction, both that a person is not to be identified merely with a particular body or a particular consciousness, and that it is logically possible for a person to exist apart from a body. It should be apparent, however, that this claim does not provide support for his conception

of a person as an entity to which both states of consciousness and corporeal characteristics are equally available.

Good reasons for rejecting Cartesianism

5.1 Although Strawson's arguments do not establish that a self is not an immaterial entity there are good reasons, in my opinion, for rejecting the view of a person as an 'immaterial substance' or 'immaterial thinking thing'. Firstly, this view gives rise to a host of unanswerable questions, e.g. 'How are immaterial entities to be individuated?', 'How are immaterial entities to be identified?', 'How do a body and non-physical mind interact?', and so on. And, secondly, in the light of current knowledge of psycho-neurophysiological correlations it seems more plausible to identify the mind with the brain. Let me, briefly, expand on these two points. With regard to the first point I shall discuss only the last of the 'unanswerable questions' I have listed, for if it were possible to explain how a body and a non-physical mind could interact the Cartesian view/^{would,} it seems to me, be a great deal less suspect.

5.2 A notable modern philosopher who has criticized Descartes' dualistic conception of a person on the ground that it is impossible to explain how a body and a non-physical mind can interact is Ryle. He says, in The Concept of Mind

" . . . the problem of how a person's mind and body influence one another is notoriously charged with theoretical difficulties. What the mind wills, the legs, arms and the tongue execute; what affects the ear and the eye has something to do with what the mind perceives; grimaces and smiles betray

the mind's moods and bodily castigations lead, it is hoped, to moral improvement. But the actual transactions between the episodes of the private history and those of the public history remain mysterious, since by definition they can belong to neither series" (p.12).

By 'episodes of the private history' Ryle means, it seems, happenings in and to a person's mind, and by 'those of the public history' he means heppenings in and to a person's body. These 'episodes' can 'by definition' belong to neither series because whereas human bodies are said to be 'in space', minds are not.

5.3 An attempt to meet Ryle's objection has been made by Lewis in a chapter entitled 'Ryle and Descartes' in 'The Elusive Mind'.¹ He says,

"The 'transactions' remain mysterious because they are non-existent, we say nothing about them, and are baffled if we try, because there are none. And why should there be? Even when we think of causal relations in the external world, we do not look for links^k (or 'transactions') between cause and effect to explain how the one leads to the other".

Lewis expands on this, saying

" . . the last thing we should try to do is to look for some link or transaction between one event and whatever it necessitates. We do not in the last resort explain causal relations, except in the sense of unfolding in greater detail the way things do in fact behave or of providing, in some of the ways suggested, some general justification for causality itself or for our confidence in it. We do not see why causal

¹ See pp.15-44, esp. pp. 26-29.

relations must be such as they are, we just accept what we find subject to the underlying assumption of consistency or continuity. We can, on the basis of what we know already, insist that certain things must be - and that others cannot be - but there is nothing beyond this, in the nature of the processes themselves, to show us why they are followed by certain others."

Lewis' argument, it is plain, has been greatly influenced by Hume's account of the relationship between causes and their effects. Hume, it will be recalled, argued that there is no necessary connection between causes and their effects on the ground that we cannot infer, through reason alone, what effect a particular cause will have.¹ 'Adam, though his rational faculties be supposed, at the very first, entirely perfect', said Hume, 'could not have inferred from the fluidity and transparency of water that it would suffocate him, or from the light and warmth of fire that it would consume him.'² And even after we have discovered, through experience, that a particular effect follows a particular cause we cannot, according to Hume, produce a 'chain of reasoning' to prove that they will always be associated. For it is always possible to imagine that a cause will have a different effect in the future. Whether Hume is correct in thinking that there is no necessary connection between causes and their effects is questionable,³ though he is, I believe, correct in thinking

¹ Hume, D., Enquiry Concerning Human Understanding, ed. L.A.Selby-Bigge (Oxford, 1902), Sections iv and vii.

² op. cit., sec. iv.

³ For there is, it seems, a sense in which the connection between a particular cause and a particular effect is a necessary one. What is wrong with saying, for example, something like 'When a match is struck a flame must result, given that such-and-such conditions are fulfilled' or 'When water is heated to 100°C at sea level, it must boil, given that such-and-such conditions are fulfilled'?

that we can give no 'ultimate reason' why a particular effect is associated with a particular cause. But let us see why he is correct, and whether in raising this point, Lewis is able to meet Ryle's objection.

5.4 Consider the question 'Why does a kettle boil when placed on a fire?' Someone might attempt to answer this question by saying something like 'The reason why a kettle boils when placed on a fire is because the heat of the fire causes the kinetic energy of the molecules to move about in a rapid fashion.' But this answer gives rise to other questions like 'Why does the heat cause the kinetic energy of the molecules to increase?' and 'Why does the rapid movement of the molecules cause the water to boil?' And an answer to each of these questions (assuming that it is possible to answer them both) will give rise to other questions, answers to which will give rise to more questions. For it will always be possible to ask 'But why is this the case?' Any attempt to explain why a particular cause leads to a particular effect will, therefore, as Lewis maintains, lead 'to a hopeless infinite regress!'

5.5 Lewis is, however, mistaken, in my opinion, in thinking that he can meet Ryle's objection by pointing out that we do not in the last resort explain why a particular cause is followed by a particular effect. For while we can give no "ultimate reason" why a particular cause is associated with a particular effect, we can, very often, offer a partial explanation of why the two are associated by providing a list of the intervening processes between them. In all cases of causal ties in the material world it is possible, in my opinion, either to provide a list of the intervening processes between causes and their effects - prior to a kettle's boiling when placed on a fire, for example there occur a number of intervening processes: there is an increase in the kinetic energy of molecules in the water, an increase in movement among the

molecules, etc,- or, in cases where there are no intervening processes, to provide mathematical equations which give 'accurate quantitative expression to what happens, where, and when.'¹ But this clearly is impossible in the case of the supposed causal relationship between physical events and mental events. We have no idea what intervening processes could occur between a particular brain state and a particular state of a non-physical mind. Nor have we any idea how to explain what happens using mathematical equations.²

5.6 Let us now turn to my second reason for rejecting the Cartesian view of a person, viz., that in the light of current knowledge of psycho-neurophysiological correlations it seems more plausible to identify the mind with the brain. A possible objection to this might be that there is no conceivable experiment which could ever prove that the mind is the brain. The most that could be established through neuro-physiological research, it might be said, is that whenever a person is in a particular mental state his brain is also in a particular state. But even if this were all that could be established (and I am not sure that it is), it would provide good evidence, in my opinion, for identifying the mind with the

¹ See Campbell, K., Body and Mind, p 37. He points out that there are no intervening processes in the case of an atomic nucleus absorbing a flying particle and then dividing.

² It should perhaps be noted that the objection that it is impossible for the Cartesian to explain how a body and a non-physical mind interact, can be avoided, by maintaining that although mental and physical events are correlated in a regular way there is no causal connection between them. On this view, which is known as the Parallelist view, it just so happens that whenever a particular mental event occurs a particular physical event also occurs, and vice versa. It is purely accidental, for example, that whenever I stub my toe I feel pain. I am doubtful, however, whether this view is any more acceptable than the interactionist view, for it seems highly improbable that such constant correlations could occur by chance.

brain. For to postulate the existence of a non-physical state which is correlated with a brain state is to complicate matters unnecessarily. By identifying mental states with brain states "Not only are superfluous entities disposed of but the subsistence of the correlations becomes explicable. Without this, these correlations would have to be just accepted as brute and inexplicable facts which, moreover, could not be fitted into the general scientific framework. Correlation 'laws' relating the intersubjectively confirmable with the non-intersubjectively confirmable would be, as Feigl expresses it, 'nomological danglers.'¹ Moreover, the view that the mind is the brain has, I believe, been shown to be free from logical objections.² A traditional objection to the mind-brain identity theory, for example, has been that while it makes sense to say of a molecular movement in the brain that it is slow or swift, straight or circular, it does not make sense to say this of an experience. The identity theorists have met this objection by arguing that although the expression 'experience' and 'brain process' do not have the same logic they may refer to the same thing. The expression 'cloud' and 'mass of tiny particles' are logically independent, yet we do not, Place has pointed out, conclude that they must refer to two separate entities. The same is, he has said, true of the expressions 'lightning' and 'motion of electric charges'.

¹ 'Editors Introduction', The Mind/Brain Identity Theory, edited by C.V.Borst, p 28.

² On this point see : Feigl, H., 'Mind-body, not a pseudo-problem'; Place, U.T., 'Is consciousness a brain process?'; Smart, J.J.C., 'Sensations and brain processes'; Armstrong, D.M., 'The nature of mind'. All of these articles are reprinted in The Mind/Brain Identity Theory.

SUMMARY

6.1 Strawson, we have seen, rejects the Cartesian view of a person because he thinks that one could have no conception of oneself as a subject of experience unless one could ascribe experiences to others, and one could not ascribe experiences to others unless they were entities to which both P-predicates and M-predicates were equally applicable. But it does not, I have said, follow from these epistemological points that persons must be essentially corporeal entities; there is nothing Strawson has said which prevents the dualist from agreeing with his view of how we acquire the idea of a subject of experience, and yet denying that persons are essentially material entities. Strawson may, I argued, want to defend his view of a person, for a disembodied person, according to him, must always think of himself as a former person 'in order to retain his idea of himself as an individual'. I hope to have shown, however, that there is no good reason for supposing this to be the case. I pointed out that since a disembodied person can, on Strawson's view of survival, observe what is going on in the universe, have certain feelings for various members of the human race, etc., it is more than likely that he will continue to broaden and deepen his experience. He might, I said, even realise some latent talent, and this might transform his whole being. But although Strawson's objection to Cartesianism is unsuccessful, there are, I have indicated, good reasons for rejecting the Cartesian view of a person. The conviction that persons are immaterial entities is not obviously false, but in the light of neurophysiological research, I have maintained, it seems more plausible to identify the mind with the brain. Moreover, in identifying the mind with the brain, we have seen, we overcome such difficulties

as explaining how immaterial persons are to be individuated and identified, and how a body and a non-physical mind can interact, etc. These reasons for rejecting Cartesianism are, it must be stressed, by no means intended to be absolutely compelling.

I have tried merely to show, as briefly as possible, why I find the Cartesian view to be a most unsatisfactory one. I hope, however, to provide more detailed support for a materialist view of a person in the following chapter.

CHAPTER 4

MATERIALISM AND PERSONAL IDENTITY

INTRODUCTION

1.1 The prime purpose of this chapter is, I have said, to reveal that the problem of personal identity, i.e., the problem of clarifying what criteria we use when we reidentify persons, is to be decided, not in favour of the memory criterion, but the bodily criterion. I hope to show that although the 'puzzle cases' generated by the problem of personal identity do, at first glance, seem to favour the memory criterion, they do not, in fact force us to concede that the memory criterion is the most important criterion of personal identity. The question of whether the problem of personal identity is to be resolved in favour of the bodily criterion or the memory criterion is, however, a difficult one, for one can, I think, in the final analysis, easily be sceptical of both of these criteria. Against those of us who think that the bodily criterion is the more fundamental of the two, for example, it can be said that since it is not inconceivable that one could wake up one morning with the same set of memories one now has but with a completely different body, the idea of a bodily transfer is not absurd, and personal identity cannot, therefore, be said to depend on bodily identity. Nevertheless, I shall argue, it is wiser to favour the bodily criterion

Strawson and Locke on Personal Identity

2.1 Towards the end of the chapter entitled 'Monads' in Individuals Strawson remarks that no attempt to solve the problem of personal identity is likely to be successful unless the thesis that a person is an entity to which

both states of consciousness and corporeal characteristics are equally ascribable is clearly understood and admitted.¹ This is so, it seems, because he thinks it would be impossible to individuate persons (and hence to reidentify them) if they were not material entities. He has argued, it will be recalled, that it would be impossible to identify others if they were Cartesian egos. And he has rejected the view that we can identify a person as a Cartesian ego via the reference to the body with which he is causally related on the ground that there is no guarantee that there are not any number of Cartesian egos attached to a single body. 'Uniqueness of the body', according to Strawson, 'does not guarantee uniqueness of the Cartesian soul.'² Now, apart from not finding this in the least convincing, Strawson is, I have already argued, himself committed to the Cartesian view.³ It seems, therefore, that his solution to the problem of personal identity would be very similar to Locke's. Locke claimed that 'the sameness of a rational being' is dependent on possession of 'the same consciousness', i.e., possession of the same set of memories.⁴

¹ p 133

² op. cit., p 101

³ For a persuasive criticism of this argument see: Kim, Chin-Tai, 'Cartesian Dualism And The Unity Of A Mind', Mind, 1971.

⁴ It should be noted that his use of the expression 'consciousness' is not entirely consistent, and that he appears to use the expression to embrace both the idea of awareness and the idea of memory. This is evident, for example, in the following statement: 'For, since consciousness always accompanies thinking, and it is that which makes every one to be what he calls self, and thereby distinguishes himself from all other thinking things: in this alone consists personal identity, i.e., the sameness of a rational being; and as far as this consciousness can be extended backwards to any, past action or thought, so far reaches the identity of that person' (Essay Concerning Human Understanding, Book 2, Ch. 27, section 9).

If the 'same consciousness' is preserved, 'whether in the same or different substances,' then, said Locke, 'the personal identity is preserved'. Thus,

"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, everyone sees he would be the same person with the prince, accountable only for the prince's action . . ." ¹

This, it would seem, is also Strawson's view. For both Locke and Strawson, therefore, saying that the person before one is the same as the person one knew before does not depend on that person's having the same body. A person's identity depends, in their view, on the set of memories which he has. Strawson, however, unlike Locke, would not hold that memory is the sole criterion of personal identity. ²

¹op.cit., section 15.

²According to Strawson, 'The criteria of personal identity are . . . multiple' (p 133)

Memory as a criterion of person identity.

3.1 That Locke was mistaken in thinking that he could define personal identity solely in terms of memory has, I think, been made abundantly clear.¹ What he failed to realise, it seems, is that memory claims require corroborative evidence for, although, as Shoemaker has argued, it is a logical truth that memory claims are usually true,² people do sometimes make false memory claims. If, therefore, a person claims to remember performing certain actions and witnessing certain events, we have no reason to believe him unless we have checked his memory claims. Ordinarily if a person X claimed to remember performing a certain action A and witnessing a certain event E, and we wanted to check these, we would have to find out whether anyone had any memory of X's performing A and witnessing E. Someone could remember X's performing A and witnessing E, however, only if X were bodily present when A was performed and E was witnessed, which shows that bodily identity must be a criterion of personal identity.

¹ By, for example, Sydney Shoemaker's 'Personal Identity and Memory' in The Journal of Philosophy, Vol.56 (1959), pp 868-882, and Bernard Williams' 'Personal Identity And Individuation' in Proceedings of the Aristotelian Society, Vol.57 (1956-57), pp 229-252. This section owes a considerable debt to both of these articles.

² One of his arguments for this is, briefly, the following. A primary criterion for determining whether a person understands the meaning of the term 'remember' is whether the confident claims of the form 'I remember such and such' that he makes are generally true. But to suppose that it is only a contingent fact, which could be otherwise, that confident perceptual and memory claims are generally true is to suppose that we have no way of telling whether a person understands the use of words like "see" and "remember", or means by them what others mean by them. . . And this is a logically absurd supposition' (Self-Knowledge and Self-Identity, pp 231-232).

3.2 In a case of 'bodily transfer' the way in which we would set about checking a person's memory claims would be slightly different, viz., because the person who claims to remember doing certain actions and witnessing certain events no longer has the same body. If, for example, a man with a cobbler's body claimed that he really was a prince and that he could remember performing certain actions as a prince, the way in which we would set about checking his claims would be by finding out, whether they fit the life-history of some one prince in the past. If we could not connect his memory claims with the life-history of some one prince then we would have no reason to believe his claims. Now, supposing that his claims fit the life pattern of, say, Prince Metternick, we would succeed in finding out that his memory claims fit Prince Metternick's life history only if we knew what Prince Metternick did, and we could know what he did only because his activities have been witnessed and recorded. His activities could not have been witnessed, however, unless he had been physically present when they took place. So once again we see that the memory criterion cannot be the sole criterion of personal identity. Further support for this is also provided by the requirement that the person making the memory claim must be reidentifiable as the same person throughout the period during which his memory claims are uttered and checked.

3.3 Some philosophers might now, in view of the above arguments, want to argue that since checking a person's memory claims involves finding out whether he was physically present when the actions he claims to remember performing occurred, that memory cannot be a criteria of personal identity at all, and that the bodily criterion is thus the sole criterion. But this, for one thing, is a self-defeating move. For in checking some person's, say X's, memory claims we have to rely on the memories of those who have witnessed the activities described by X. We have to rely on their having remembered seeing X performing certain actions and witnessing certain events. Furthermore, the fact that we find bodily transfer cases puzzling in itself suggests that memory is a criterion of personal identity. The fact that we are in doubt as to whether we should decide in favour of the bodily criterion or the memory criterion would appear to indicate that neither criterion can be abandoned completely. The one criterion may, of course, be more fundamental than the other. Strawson, I have already indicated, would hold that the memory criterion is the more fundamental of the two, though he would not hold, as Locke does, that it is the sole criterion. And, indeed, the puzzle cases do, at first sight, appear to favour the memory criterion. What we must consider next, therefore, is whether the so-called 'bodily transfer' cases do force us to hold that a person is to be identified primarily by the memories he has, and that the memory criterion is thus more fundamental than the bodily criterion.

3.4. To return to the case of the cobbler and the prince, if we suppose that the memory claims of the man who physically seems to be a cobbler do fit the life history of Prince Metternick, and if we suppose also that some of his memory claims that cannot be checked throw light on certain previously unexplained events, are we to say that the cobbler has become Prince Metternick, that Prince Metternick has come to life once more in a cobbler's body? Bernard Williams certainly does not think we are forced into accepting this conclusion.¹ He considered a similar case in which a man called Charles makes certain memory claims that fit the pattern of Guy Fawke's life, and he rejects the view that we are forced to accept that Charles is identical with Guy Fawkes on the ground that 'it is not logically impossible that two different persons should claim to remember being this man and this is the most we can get.'² He attempts to provide additional support for holding that Charles is not to be said to be identical with Guy Fawkes by advancing the following argument.

'If it is logically possible that Charles should undergo the changes described, then it is logically possible that some other man should simultaneously undergo the same changes; e.g., that both Charles and his brother Robert should be found in this condition. What should we say in that case? They

¹ See his 'Personal Identity and Individuation', PAS, Vol.57 (1956-57), esp. pp.238-241.

² p.238.

cannot both be Guy Fawkes; if they were, Guy Fawkes would be in two places at once, which is absurd. Moreover, if they were both identical with Guy Fawkes, they would be identical with each other, which is also absurd. Hence we could not say that they were both identical with Guy Fawkes. We might instead say that one of them was identical with Guy Fawkes, and that the other was just like him, but this would be an utterly vacuous manoeuvre, since there would be *ex hypothesi* no principle determining which description was to apply to which. 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. 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?'¹

Williams concedes that this argument does not prove that we cannot speak of identity in the simpler case where just Charles' memory claims fit the pattern of Guy Fawkes' life, but he maintains that it does show that to speak of identity in the simpler case would be quite vacuous. He attempts to clarify this point by referring to a distinction between identity and exact similarity. He distinguishes the two terms by maintaining that there is an obvious difference in saying that two men live in the same house, and that they live in exactly similar houses. He then goes on to say that the distinction can only be drawn in the case of material objects. In the case of character the distinction cannot be drawn, for in saying that A and B have the same character one is simply saying that they have exactly similar characters. In the

¹ op. cit., pp. 238-239.

case of memories it appears that it is not even possible to say that A and B have exactly similar memories. For, to return to the Charles-Fawkes case, we cannot say that Charles 'has the same memories as Guy Fawkes, as this is to imply, what we want to deny, that he really is Guy Fawkes; nor can we say that the memory claims he makes are the same as those made by Guy Fawkes, as we have little idea of what claims Fawkes in fact made, or indeed of how much he at various times remembered. All we actually know is that Charles' claims fit Fawkes' life.' Williams thus concludes that 'the omission of the body takes away all content from the idea of personal identity.'¹

3.5 What are we to make of Williams' arguments?

He has not established, in my opinion, that we cannot speak of identity in the case of memories, that we cannot speak of two persons having the same memories. That he has not established this can be seen by considering the following imaginary case.² Two men, A and B, have been simultaneously operated on for brain tumours. At the end of the operation, which involves a technique whereby a person's brain is extracted from his head and then operated on, A's brain is mistakenly put into B's head, and B's brain is put into A's head. The person with A's body and B's brain immediately dies,

¹ op.cit., p.241.

² This example is borrowed from Shoemaker, See his Self-Knowledge and Self-Identity, pp.23-24.

but the other person with B's body and A's brain survives. Upon regaining consciousness he (i.e., the person with B's body and A's brain) displays great shock and surprise at seeing his body. He recognizes A's wife and family (whom B had never met) and can describe in detail events in A's life. He has no knowledge at all of B's life. Are we to say in this case that the person with B's body does not really have A's memories, that his memories and A's memories are not one and the same? Surely not, for if we know that the person with B's body has A's brain then this would explain how it is that he has A's memories. We are justified in holding that he has A's memories, it seems, simply on the ground that we know that he has A's brain. It is, of course, possible that he might continue to remember in detail events in B's life, and have no knowledge of A's life at all. If this were the case we would then be forced to conclude that there is no close causal relationship between the states of a man's brain and his psychological states, and that we were mistaken in thinking that there was. In the light of present scientific knowledge, however, it is not implausible to hold there is such a close causal relationship. But if it is not implausible to suppose that the person with B's body and A's brain will have A's memories - it is, I believe, highly plausible - and if, as I have assumed, the idea of a brain transplant is not an inconceivable one, then clearly Williams is mistaken in thinking that we cannot speak of two men having the same memories. Williams' reduplication argument (i.e.,

the argument in which both Charles and his brother Robert claim to remember being GuyFawkes), then, does not show that to speak of identity in the simpler case where just Charles' memory claims fit Guy Fawkes' life, would be utterly vacuous. But it does prove something, as I shall shortly reveal.

3.6 Now, although I have argued that Williams is mistaken in thinking that we cannot speak of two men having the same memories - it is, I have argued, possible to hold that the person with B's body has the same memories A had - I am not at all sure that the two men can be said to be identical. Even after one has considered the puzzle cases there are, in my opinion, grounds for claiming that personal identity depends on bodily continuity. These grounds will be advanced in the following section, in which I shall attempt to show that it is wiser to favour the bodily criterion rather than the memory criterion.

A prejudice in favour of bodily identity.

4.1 Most philosophers who think that the puzzle cases favour the memory criterion assume, in my opinion, that we have a clear concept of a psychical entity which can leave one body and enter another. Penelhum is, I think, correct, when he says :

"The reason that we all feel some degree of compulsion towards accepting the bodily-transfer solution is that dualist pre-conceptions intrude themselves when we investigate the stories. It is taken for granted that we have an independently clear concept, with recognised criteria of identity, of a soul,

spirit, or mind, which can be thought of as having a purely contingent relationship to the body, which it may abandon in favour of another body."¹

Only bodily transfer cases which are explained in terms of brain transplants, in my opinion, give content to the idea of a person leaving one body and entering another. For here we are able to make an independent identification of an entity which can be taken out of one body and put into another. Bernard Williams' arguments in the previous sections do, therefore, reveal that, unless an individual has undergone a brain transplant, we are not forced to admit that an individual who has A's memories and B's body has changed bodies, for all we need say is that B has somehow acquired A's memories. But even if there is a logical possibility of a bodily transfer taking place, there are, I believe, good grounds for holding that a person who undergoes a bodily transfer cannot be said to be the same person he used to be.

4.2 Consider the following imaginary example.

Two men, A and B, are involved in a motor-car accident. Both are seriously injured and taken to hospital. A, it is discovered, has suffered irreparable brain damage. B, on the other hand, has received only minor internal injuries, but his body has been badly burnt and mutilated - he has lost an ear and an eye, and half of his jaw has been severed from his face. As a result of being so terribly disfigured he is not even recognised by one of the doctors who is a close friend of his. The surgeons at the hospital to which the

¹ Penelhum, T., 'Personal Identity', The Encyclopaedia of Philosophy, ed. by Paul Edwards, Vol.6., p.106.

two men have been taken, decide, therefore, that since A will be a vegetable if he lives, it would be a good idea to extract B's brain from his badly burnt and mutilated body and put it into A's body, which has suffered only minor external injuries. The surgeons are convinced that both B and his wife will be delighted when they learn of their decision to replace B's mutilated body with A's body, especially as A and B have such similar physiques.¹ They also believe that, in view of the fact that A and B have roughly similar physiques, B's personality will not be affected in any way. But would B's wife be delighted at the result of the operation? And would B's personality be unaffected? These are the questions which we must now consider. If our answers to them are negative then this will show that B would not be the same person he was before the accident.

4.3 Let us first consider what B's wife's reaction to 'her husband' would be. I suggest that her initial reaction might well be to refuse to accept that her husband's brain has been transplanted into another man's body, that the man with her husband's memories really is her husband. This, I believe, would be an extremely natural reaction. After some persuasion, however, she might reluctantly accept that the man who claims to be her husband does have his brain, but continue to deny that he is the same person he used to be. She might claim that, because of his having acquired a different body, and a different voice, she feels like a complete stranger in his presence, and that she finds responding to his affection exceptionally difficult. In order to see this more clearly

¹ It is, I think, permissible to speak of the man with A's body and B's brain, in a loose sense, as B, in that it would surely be absurd to deny that B had survived. But I shall expand on this point in due course.

let us alter our example slightly, and suppose that A's body i.e., the body into which B's brain has been transplanted, is a female body or the body of a friend of both B's and his wife. B's wife would obviously feel considerably uneasy about calling either a person with B's brain and a female body, or a person with B's brain and the body of a friend of hers, her husband. Her relationship with either of these persons could never be the same as the relationship she had with her husband before his accident. It would seem, therefore, that personal identity is very much dependent on bodily identity.

4.4 But let us next consider whether B's personality would be drastically affected by his change of body.¹ Owing to the way in which the example has been set up it is not, in my opinion, easy to assess whether B's personality would be affected or not. A, it was said, had a roughly similar physique to B, thus diminishing the possibility of B's personality being affected. If, however, we again alter our example slightly, and suppose that B's body was very different to A's, then it is obvious that B's personality would be seriously affected. Imagine, for example, how B's personality would be affected if, after possessing a very muscular and athletic type of body, he acquired a body which was very small and frail, or, worse still, he acquired a female body. So again we can conclude that personal identity is dependent on bodily identity.

4.5 A possible objection to the view that personal identity depends on bodily identity which might be raised is

¹Most people would, I assume, agree that only if a person undergoes a drastic personality change can we say that he is not the same person he used to be. For it is I think true that most of us undergo slight changes in our personality during our lives, and yet still remain basically the same people.

the following. Even if B's brain is not transplanted into A's body, it might be argued, B's personality will still be affected, owing to the extent to which his body has been mutilated. And if this is so then possession of the same body is not a necessary condition of personal identity. Now philosophers who might raise this objection are entitled, in my opinion, to assume that if a person is involved in an accident and his body is badly mutilated, then his personality could be seriously affected; for he could easily develop any number of complexes about his appearance, and so become very depressed and neurotic, very bitter and resentful, etc.; and this, in turn, could affect his relationship with, say, his wife and family, so that they would have a further reason for not regarding him as the same person he was before his accident. What is not clear, however, is whether the philosophers who might raise this objection are entitled to assume that B has the same body after the accident. The reason why this is not clear is, not because it seems paradoxical to say that something has changed and yet it is still the same, but, rather, because of the extent to which B's body has changed. B's body has changed to such an extent that his personality is bound to be seriously affected by the change. Sameness and change, it should be noted, are not, as Hume thought, incompatible. Most of our substantive concepts are, as Penelhum has pointed out, designed to incorporate changes.¹ But they do not, as has also pointed out, permit all kinds of change indiscriminately. How much change is permitted depends

¹ op.cit., pp. 99-100.

according to Penelhum, on the concept in question. 'A man', he says, 'can change in more ways before he is destroyed than a chair can. To know what alterations are and are not allowed is to know, among other things, what the criteria of identity are for the class of entities grouped under the concept in question.' There is, however, an important difference between persons and other material entities, in that it is not always just the alterations in a person's appearance which incline us to say that he is not the same person he used to be.

4.6 Another possible objection to my defence of the view that personal identity depends on bodily identity which might be raised is that extreme cases only reveal that a man cannot change his body for one that is vastly different to his own, and not that personal identity depends on possession of the same body. If a man were to acquire a body which was very similar to the one he originally possessed, then, it might be argued, there seems to be no reason for supposing that his personality will be affected in any way. I am not at all sure that this is true. But even if this is granted, we could, it seems to me, still defend our view. For although the person who has changed bodies does not undergo a change in personality, it is unlikely that the change will not have an affect on others, or at least those close to him, and that they will still regard him as the same person he was before he changed bodies. Imagine, for example, how the wife of a man who had changed bodies would feel about making love to 'her husband'. But even if all that I have said is mistaken, the consideration of extreme cases does reveal (i) that it is a necessary condition of personal identity that a

person possesses a body which is very similar to the one he used to have; and (ii) that there is a severe limitation to the number of bodily transfer cases which can be constructed, and that bodily identity will, therefore, be a sufficient criterion of personal identity in most cases.

Identity and Survival

5.1 Having attempted to demonstrate that bodily continuity is a necessary condition of personal identity, I wish now to maintain that it does not follow from this that bodily continuity is a necessary condition of survival. I argued, in the previous section, that a person could not be called the same person if, as a result of acquiring a new body, his personality was drastically changed, or other people would not accept that he was the same person. But it does not follow from this, in my opinion, that a person's change of body is equivalent to his death. Few people would, I believe, deny that if a person's brain, say A's brain, were transplanted into another person's body, and the resultant person had A's memories, then A had survived. How many people who were in the position of a person who was going to have his brain transplanted into another person's body would willingly consent to the torture of the other person's body on completion of the transplant? Some people, however, might accept that survival is not dependent on bodily continuity, and yet deny that personal identity is, as I have claimed, dependent on bodily

continuity. Survival, they might argue, implies identity, But it can, I think, be shown that admitting that a person can survive a change of body does not force one to hold (i) that he is the same person he used to be, and (ii) that bodily identity is not, therefore, a necessary condition of personal identity.

5.2 That survival does not presuppose the retaining of identity has, I think, been ably demonstrated by Derek Parfit.¹ He discusses an example which is very similar to Williams' example in which two brothers Charles and Robert both claim to remember being Guy Fawkes. The example he discusses is one in which his brain is bisected, and the one hemisphere is put into one person's body, the other into another person's body. Both of the resulting people have his character and memories. In response to the question 'What happens to me?' Parfit offers three possible answers. They are (1) I do not survive; (2) I survive as one of the two people; (3) I survive as both. All three answers are rejected by him. He rejects the first answer as follows. We agreed that he could survive if his brain were successfully transplanted into another person's body. And people have actually survived with half their brains destroyed, so he could survive if half of his brain were successfully transplanted and the other half destroyed. But if this is so, how could he not survive if the other half

¹ 'Personal Identity', The Philosophical Review, vol.80, No.1 (January, 1971).

were also successfully transplanted? 'How could a double success be a failure?' The reason that Parfit rejects the second answer is that there is nothing which can make him the one person rather than the other. Each of the resulting persons is, to start with, exactly similar. So how could he survive as only one of the two people? Finally, he rejects the third answer on the ground that although the resulting persons will, initially, be exactly similar, they will not continue to be. Each person will develop differently according to the influences he is exposed to. 'They could live at opposite ends of the earth. (If they later met, they might even fail to recognize each other). It would become intolerable to deny that they were different people.' So although Parfit thinks it is a mistake to deny that he survives, he thinks it is also a mistake to hold that he survives as either one of the two people or as both of them. We should, according to him, hold that he survives as two different people without implying that he is those people. The sense in which one man survives as two is, he says, in the sense that 'the two resulting people are his later selves. And they can each refer to him as "my past self". (They can share a past self without being the same self as each other).'

5.3 Parfit, it should be noted, does not maintain that we cannot speak of identity when a person's brain is transplanted into another person's body. He argues, against Williams, that 'psychological continuity' does provide a criterion of identity.¹ He holds that it provides a ground for speaking

¹ He defines 'psychological continuity' in terms of quasi-memory.

of identity, however, only if the relation is one-one. If psychological continuity always took a 'one-many or branching form' then we could not speak of identity at all. He agrees with Williams, therefore, that identity is a one-one relation, but disagrees that Williams' reduplication argument proves that since psychological continuity is not logically one-one, it cannot provide a criterion of identity. If one man is psychologically continuous with Guy Fawkes we cannot, according to Parfit, deny that he is Guy Fawkes. To do so would be to refuse to obey Williams' principle that an important judgment should be asserted and denied only on importantly different grounds. In the case of two men being psychologically continuous with Guy Fawkes we have an important ground for saying that Guy Fawkes has survived as two people. Parfit concludes that 'Even if psychological continuity is neither logically, nor always in fact, one-one, it can provide a criterion of identity. For this can appeal to the relation of non-branching psychological continuity, which is logically one-one.'

5.4 But enough of Parfit. He has, I think, made it abundantly clear that survival does not presuppose the retaining of identity. For if one person can survive as two people, and if the resulting people are not identical with each other - they may, to start with, be exactly similar,

but clearly they will not continue to be - then they cannot be said to be identical with the original person. For what could make one of them identical with him rather than the other? So survival does not imply the retaining of identity.

Williams' Puzzle

6.1 having discussed the problem of personal identity at some length we are, I think, in a position to consider the intriguing puzzle created by Bernard Williams in his article 'The Self and The Future'.¹ Williams presents us with two imaginary cases, the first of which seems very much to show that personal identity is dependent on memory, and the second of which seems very much to show that personal identity is dependent on bodily identity. This is particularly puzzling in view of the fact that the two imaginary cases are really just the same case differently presented.

6.2 The first imaginary case he presents us with is the following. Two men, A and B, are to enter a machine which will extract the information contained in each of their brains, and replace it in the other's brain. The 'A-body-person' will thus emerge with the information originally contained in B's brain, and the 'B-body-person' will emerge with the information originally contained in A's brain. Before entering the machine A and B are told that one of the two resultant persons is going after the experiment to be given \$100,000, while the other is going to be tortured. They are

¹ The Philosophical Review, vol.79, No.2. (April 1970)

then asked to choose, on selfish grounds, what treatment should be dealt out to each of the resultant persons emerging from the experiment. A chooses that the 'B-body-person' should get the money and the 'A-body-person' the pain, and B chooses conversely. The experimenter, who cannot act in accordance with both sets of preferences, decides to act in accordance with B's wishes, and inflict pain on the 'B-body-person', and give the money to the 'A-body-person'. After the experiment the 'B-body-person' will, according to Williams, not only complain about the unpleasant treatment, but he will complain (since he has A's memories) that this was not the outcome he chose, for he chose that the 'B-body-person' should be well treated. The 'A-body-person', on the other hand, will thank the experimenter profusely for giving him the \$100,000, and for deciding to act in the way that he, B, so wisely chose. These facts, Williams concludes, 'make a strong case for saying that the experimenter has brought it about that B did in the outcome get what he wanted and A did not. It is therefore a strong case for saying that the 'B-body-person' really is A, and the 'A-body-person' really is B; and therefore for saying that the process of the experiment really is that of changing bodies.'

6.3 Williams attempts to reinforce the view that the imaginary example demonstrates that there is a strong case for saying that a change of bodies has taken place, by considering some alternative choices which A and B could make.

I shall mention what he has to say about just one of them. Suppose, he says, that A chooses that the 'A-body-person' should get the money, and the 'B-body-person' the pain, and B chooses conversely. And suppose that the experimenter decides to comply with A's wishes, and give the 'A-body-person' the money, and the 'B-body-person' the unpleasant treatment. After the experiment the 'B-body-person' will acknowledge, according to Williams, that he got what he chose (since he has A's memories), but he will not like what is happening to him. The 'A-body-person', on the other hand, will claim that he did not get what he chose and that this was most fortunate for him. So once more, says Williams, it looks as though the 'A-body-person' and the 'B-body-person' are, respectively, B and A.

6.4 The second imaginary case Williams presents us with is the following. Someone who has power over him tells him that he is going to be tortured the next day. This person also tells him that shortly before the torture he will not remember being told he was going to be tortured. But this, says Williams, does not help to cheer him up, for he knows that he can forget things, and that there is such a thing as being tortured unexpectedly because he had forgotten or being made to forget a prediction of torture. He is then told that he will not remember any of the things he is now able to remember. But this also does not help to cheer him up, for he knows that he can be in an amnesiac state and also in great pain. Finally, he is told that he will have a completely different set of

memories. But this does not cheer him up either, for he knows that there is a possibility of his going mad, and thinking that he is someone else; and being told that something like this was going to happen to him would not reduce the fear of being told that he was going to be tortured. Williams concludes that when he reflects on the grounds he has for fearing that he will be tortured he is aware that his fear rests on the seemingly sound principle that his undergoing physical pain in the future is not excluded by any psychological state he may be in at the time. This, then, appears to support the view that personal identity is dependent on bodily identity.

6.5 How are we to resolve Williams' puzzle ?

I think we can resolve his puzzle by demonstrating that his description of the outcome of the experiment, in which A's and B's memories are swapped, does not make a strong case for saying that the process of the experiment really is that of changing bodies. Williams' description of the outcome of the experiment can, it seems to me, easily be seen as incomplete. Let us first take his description of the outcome of the experiment where A chooses that the 'B-body-person' should get the money and the 'A-body-person' the pain, and B chooses conversely, and the experimenter decides to act in accordance with B's wishes. According to Williams, the 'B-body-person' will complain (since he has A's memories) that this was not the outcome he chose, whereas the 'A-body-person' will express satisfaction at the fact that the

experimenter decided to act in the way that he, B, so wisely chose. But could it not be said that both the 'B-body-person' and the 'A-body-person' were mistaken in their thinking? Could we not say to the 'B-body-person' something like, 'You did really get what you chose, and your choice was an unwise one. The reason why you think that you didn't get what you chose is because you have A's memories'? And could we not say to the 'A-body-person' something like, 'You didn't really get what you chose, and you're lucky that you didn't. The reason why you think you got what you chose, and what you wanted, is because you have B's memories'? There is, it seems to me, nothing preventing us from holding that both the 'B-body-person' and the 'A-body-person' are mistaken in their thinking. This, indeed, is easily explained by the fact/^{that}they have each other's memories. So there is, after all, no strong case for saying that A and B have exchanged bodies.

6. 6 Let us next consider Williams' description of the outcome of/^{the}experiment where A chooses that the 'A-body-person' should get the money, and the 'B-body-person' the pain, and B chooses conversely, and the experimenter decides to act in accordance with A's wishes. According to Williams, the 'B-body-person' will acknowledge that he got what he chose, and that his choice was an unwise one, whereas the 'A-body-person' will claim that he did not get what he chose, and that he was very lucky that he didn't. But could it not again be said that both the 'B-body-person' and the 'A-body-

person' were mistaken in their thinking. Could we not say to the 'B-body-person' something like, 'You didn't really get what you chose. The reason why you think you did, and that your choice was an unwise one, is because you have A's memories'? And could we not say to the 'A-body-person' something like, 'You did really get what you chose. The reason why you think that you didn't, and that you're lucky you didn't, is because you have B's memories'? It is not at all clear, therefore, that it is rational to make a choice as if one's identity depended on one's memories.

6. 7 A possible objection to what I have said above must now be met. It might be argued that if I am prepared to concede that the 'A-body-person' and the 'B-body-person' would not be mistaken in their thinking if B's brain were transplanted into A's body, and A's brain were transplanted into B's body, then I cannot, if I am going to be consistent in my views, deny that they will be correct in their thinking if they have just swapped memories. I think we can reply to this objection by saying, simply, that there is an important difference between extracting information from my brain and putting it into another person's body, and extracting my brain and putting it into another person's body. If my brain is extracted from my head I shall cease to exist (as a material being, at any rate), unless my brain is kept alive by, say, being put into another person's head. But if my memories are extracted from my brain I shall not

cease to exist. I shall still fear being injured, for example, even if I am told that when I am injured I will have a completely different set of memories.

6.8 Finally, it should be noted that Williams does make a tentative attempt to resolve the puzzle he has created. According to him, the arguments 'which suggested that A should identify himself with the B-body-person, turned on the extreme neatness of the situation in satisfying, if any could, the description of "changing bodies". But this neatness is basically artificial; it is the product of the will of the experimenter to produce a situation which would naturally elicit, with minimum hesitation, that description.'

Williams admits that we could say of the experiment that it involved a changing of bodies' if we had some model of ghostly persons in bodies, which were in some sense actually moved around by certain procedures', but he denies that we can 'seriously use such a model'. (It should be apparent from what has been said in section 4.1 that I am in full agreement with Williams on this point). He concludes that we should thus favour the arguments of the second imaginary case. 'The principle that one's fears can extend to future pain whatever psychological changes precede it seems positively straightforward. Perhaps, indeed, it is not; but we need to be shown what is wrong with it. Until we are shown what is wrong with it, we should perhaps decide that if we were the person A then, if we were to decide selfishly, we should pass the pain to the B-body-person'.

SUMMARY

7.1 I hope to have said enough by now to have shown that Strawson is mistaken in thinking that the memory criterion is more fundamental than the bodily criterion. I argued, it will be recalled, that Strawson would, like Locke, regard the memory criterion as more important than the bodily criterion, but that, unlike Locke, he would not think of it as the sole criterion of personal identity. I then argued, after showing that Locke was mistaken in thinking that memory was the sole criterion of personal identity, that the 'bodily transfer' cases do not force us to hold that the memory criterion is the more fundamental of the two. I began my defence of the bodily criterion by pointing out, first, that we have no clear concept of a spirit or mind which can leave one body and enter another, and, second, that even if there is a logical possibility of a bodily transfer taking place, there are still good grounds for holding that a person who has the same memories as another is not identical with him. It is more than likely, I maintained, that a person who acquired a new body would not be called the same person, for two reasons, viz., because his personality would probably be seriously affected, and because those who knew him would probably not accept that he was the same person he used to be. I pointed out, however, that even if my arguments for this were mistaken they did at least establish (i) that it is a necessary condition of personal identity that a person possesses a body which is either the same or very similar to the one he used to have; and (ii) that there is thus a severe limitation to the number of bodily transfer cases which can be constructed; and (iii) that bodily identity will be a sufficient criterion of personal identity in most cases. Finally, after showing

that bodily continuity is not a necessary condition of survival and that survival does not imply identity, I presented Williams' puzzle in one further attempt to reveal that personal identity is dependent on bodily identity, and thus strengthen the view that a person is a material entity.

CONCLUSION

1.1 My chief aims in this thesis have been to establish (i) that a person is an entity which has, or is the owner of, psychological or mental attributes; and (ii) that there are good grounds for holding that the entity which has these attributes is a material entity, and not, as the Cartesian thinks, an entity consisting of two logically distinct parts, viz., a mind and a body. I have endeavoured to establish these points as follows.

1.2 In chapter one I tried to refute the view that there are no owners or subjects of psychological attributes. Wittgenstein, we saw, rejected the view that the word 'I' in a sentence like 'I am in pain' refers to a possessor of experiences on the ground that there is no question of one's having to identify anything when the person who is in pain is oneself. But what he failed to realise is that we say such things as 'I am in pain', 'I have a toothache', etc., primarily for the sake of others. A possible reply to this might be that although the 'I' refers for the hearer it certainly does not refer to the same thing to which it refers for the speaker. For the 'I', it might be said, refers, for the speaker, to a particular 'bundle' of experiences, viz., the bundle which is attached to the body with such-and-such characteristics. The trouble with holding that a person is 'a bundle or collection of different perceptions', we have seen, is, first, that it is impossible to explain how the various perceptions are connected, and, second, that one is forced into the absurd position of conceding that when one

goes to sleep one ceases to exist. The only way in which these difficulties can be avoided, I maintained, is by abandoning the view that a self is a bundle of experiences, and holding either that all of one's experiences belong to a pure ego, or that they are all causally dependent on the states of one's body. The bundle theorists' objection to the view that a self is a subject of experience, it will be recalled, also gives rise to a serious difficulty, for if, as they assume in presenting their evidence for the denial that a self is a subject, introspection is analogous to perception, it would seem that they are forced to concede that a self is a subject of experiences.

1.3 In chapter two I discussed Strawson's conception of a person as an entity to which both material body predicates and psychological predicates are equally applicable. I argued that although his claim that a person is 'a type of entity such that both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics . . . are equally applicable to a single individual of that single type' is compatible with his claim that the concept of a person is 'logically primitive', his conception of a person is really very similar to Descartes'. That this is so is, I maintained, revealed by his use of the expression 'a subject of experience' and by what he has to say about surviving the death of one's body, etc. Having argued that Strawson's conception of a person is really very similar to Descartes' I tried to show that his claim that behaviour criteria are logically adequate for the ascription of psychological predicates does not conflict with his view

of a person. It is not the case, I argued, that it is impossible to make sense of the claim that behaviour criteria are logically adequate etc. unless this amounts to the claim that behavioural evidence entails the truth of the ascription of P-predicates, or the claim that behavioural evidence provides inductive support for the truth of the ascription of P-predicates. For the sense in which behavioural criteria are logically adequate for the ascription of P-predicates is in the sense that they provide necessarily good evidence for the ascription of P-predicates; they provide necessarily good evidence, it was maintained, in virtue of the fact that it is a necessary condition of our possessing psychological concepts that we grasp the connexion between mental phenomena and behavioural facts.

1.4 In chapter three I tried to show that while Strawson has not established that a self is not an immaterial entity there are, nevertheless, good reasons for rejecting the view of a self as an 'immaterial substance'. It is a mistake, I argued, to conclude from the facts that it is a necessary condition of ascribing experiences to oneself that one can ascribe them to others, and that it is a necessary condition of ascribing experiences to others that they are entities to which corporeal characteristics are also ascribable, that persons are material entities. For all that these facts prove is that persons (or at least some persons) must be entities to which M-predicates are also ascribable, and this is not inconsistent with Cartesianism. Witness Descartes' remark: 'I can draw the certain conclusion that my body, or rather my entire self, in so far as I am composed of body and mind, can receive various pleasant or unpleasant contacts from surrounding bodies'.¹ The Cartesian does, of course, think that the

¹op. cit.,

essential part of a person is his mind. Strawson may, therefore, I have suggested, want to defend his position, for a disembodied person, in his opinion, could think of himself only as a former person. But his claim that a disembodied person could think of himself only as a former person is not, I have argued, well-founded. Having thus defended the Cartesian against Strawson's attack I concluded the chapter by arguing that, in the light of neurophysiological research, and in view of the fact that the Cartesian view gives rise to a host of unanswerable questions, the materialist view of a person is by far the more plausible of the two views.

1. 5 Finally, in chapter four I tried to provide further support for the view that persons form a class of material bodies, by arguing that the problem of personal identity is to be decided in favour of the bodily criterion rather than the memory criterion. My chief reasons for this were : (i) that we have no clear concept of a psychical entity which can leave one body and enter another; (ii) that only those puzzle cases in which an individual has undergone a brain transplant force us to admit that he has changed bodies; and (iii) that even if there is a logical possibility of a bodily transfer taking place, there are good reasons for holding that a person who undergoes a bodily transfer cannot be said to be the same person he used to be. A person who acquired a new body would not be called the same person, I suggested, firstly, because his personality would probably be seriously affected by the change, and, secondly, because even if his personality was not in any way affected those who knew him would probably not accept that he was the same

person. My discussion of Williams' puzzle has, I hope, also shown that it is wiser to favour the bodily criterion rather than the memory criterion.

1. 6 Is a person then to be identified solely by his body? It should be apparent from what I have said in the previous chapter that a person cannot be identified solely by his body. The brain, I have suggested, is also an essential part of a person. We should, therefore, identify a person, in my view, not only by his body but by his brain as well. This criterion of personal identity is, it seems, fairly similar to Wiggins'. He, apparently, thinks that " 'person' is used to pick out a spatio-temporally continuous material object which is composed of the brain and the rest of the body but not identical with either or both of them."¹ He does not think that a person is identical with his brain and the rest of his body because the loss or replacement of the rest of the body 'need not terminate the existence of the person.'² Whether a person who acquired a new body or who lost his body could be called the same person is, I have tried to show, not as obvious as many have thought. If, however, I am proved wrong on this point I would adopt Wiggins' criterion of personal identity.

¹ Smart, B., 'How can Persons be Ascribed M-predicates?', Mind, 86 (1977) pp. 49-66.

² ibid.

BIBLIOGRAPHY

In this bibliography are listed the books and articles I have referred to in the text, and also a number of other relevant works which I discovered at too late a stage to make use of.

1. Armstrong, D.M., 'The nature of mind', The Mind/Brain Identity Theory, ed. C.V.Borst (Macmillan, London, 1970).
2. Ayer, A.J., Language Truth and Logic
3. Ayer, A.J., The Concept of a Person and Other Essays (Macmillan, London, 1964).
4. Berkeley, G., Works of Berkeley, ed. G.Sampson (London, 1897)
5. Borst, C.V., (ed.), The Mind/Brain Identity Theory (Macmillan, London, 1970).
6. Campbell, K., Body And Mind (Macmillan, London, 1970).
7. Chisholm, R., 'On the Observability of the Self', Philosophy and Phenomenological Research, 30 (1969) pp.7-21.
8. Clarke, D.S., 'A Defence Of the No-Ownership Theory', Mind, 81 (1972).
9. Coburn, R.C., 'Bodily Continuity and Personal Identity', Analysis, 20 (1960) pp.117-20.
10. Coburn, R.C., 'Persons and Psychological Concepts', American Philosophical Quarterly, 4 (1967) pp.208-221.
11. Cook, J.W., 'Human Beings', Studies in the Philosophy of Wittgenstein, ed. P.Winch (Routledge & Kegan Paul, London, 1969) pp.117-51.
12. Descartes, R., Discourse On Method and the Meditations, trans. F.E. Sutcliffe (Penguin Books).
13. Feigl, H., 'Mind-body, not a pseudo-problem', Dimensions of Mind, ed. Sidney Hook (New York University Press 1960).
14. Flew, A.G.N., 'Locke and the Problem of Personal Identity', Philosophy, 26 (1951) pp.53-68.
15. Gill, J.H., "On 'I'", Mind, 80 (1971) pp.229-240.
16. Hartnack, J., 'The Metaphysics of the "I"', Philosophical Quarterly, 22 (1972) pp 248-54.
17. Hume, D., A Treatise of Human Nature, ed. L.A.Selby-Bigge (O.U.P., Oxford, 1888).

18. Hume, D., Enquiry Concerning Human Understanding, ed.L.A. Selby-bigge (Oxford,1902).
19. Jones, J.R., 'How Do I Know Who I am?', Aristotelian Society Supplementary Volume, 41 (1967) pp.1-18.
20. Jones, O.R.(ed.), The Private Language Argument (Macmillan, London,1971).
21. Kant, I., Critique of Pure Reason, trans N.K.Smith(Macmillian, London,1933).
22. Kim, C.-T., 'Cartesian Dualism and the Unity of the Mind', Mind,80 (1971) pp.337-53.
23. Lewis, H.D., 'Mind and Body - Some Observations on Mr Strawson's Views', Proceedings Aristotelian Society, 63 (1962-3) pp.1-22.
24. Lewis, H.D., The Elusive Mind (Allen & Unwin, London,1969).
25. Locke, J., Essay Concerning Human Understanding, ed. A.C. Fraser (Oxford, 1894).
26. Long, D., 'Philosophical Concept of a Human Body', Philosophical Review, 73 (1964) pp.321-37.
27. Mackie, J.L., Problems From Locke (Oxford University Press, London,1976).
28. Martin, C.B., 'People', Contemporary Philosophy in Australia, ed. Robert Brown and C.D. Rollins (George Allen & Unwin Ltd., London, 1969).
29. Moore, G.E., 'Wittgenstein's Lectures in 1930-33', Mind, 63 (1954) pp.1-15,289-316, and 64(1955) pp.1-27.
30. Moulder, J., 'In Defence of Immaterial Persons', Philosophical Papers, May 1972, pp.38-55.
31. Odegard, D., 'Persons and Bodies', Philosophy and Phenomenological Research, 31(1970) pp.225-42.
32. Parfit, D., 'Personal Identity', Philosophical Review,80 (1971) pp.3-27.
33. Pears, D.F., 'Hume on Personal Identity', David Hume : A Symposium, ed D.F. Pears (Macmillan,London,1966).
34. Pears, D.F., 'Critical Study : Individuals', Philosophical Quarterly 11 (1961) pp.172-185.
35. Penelhum, T., 'Hume on Personal Identity', Philosophical Review, 64 (1955) pp.571-89.
36. Penelhum, T., 'Personal Identity', Encyclopedia of Philosophy, ed. P.Edwards(The Macmillan Co.of New York and Free Press, New York, 1967).
37. Penelhum,T., 'Personal Identity, Memory and Survival', Journal of Philosophy,56 (1959) pp.882-903.

38. Penelhum, T., Survival and Disembodied Existence (Routledge & Kegan Paul, London, 1970).
39. Perry, J. (ed.), Personal Identity (University of California Press, Berkeley and Los Angeles, 1975)
40. Place, U.T., 'Is consciousness a brain process?' The British Journal of Psychology, XLVII (1956).
41. Quinton, A.M. 'The Soul', Journal of Philosophy, 59 (1962) pp.393-409.
42. Ramsay, I.T., 'The Systematic Elusiveness of "I"', Philosophical Quarterly, 5 (1955) pp.193-204.
43. Rollins, C.D., 'Personal Predicates', Philosophical Quarterly, 10 (1960) pp.1-11.
44. Russell, B., The Analysis of Mind (Allen & Unwin, London, 1921).
45. Ryle, G., The Concept of Mind, Hutchinson University Library (Hutchinson, London, 1949).
46. Shoemaker, S., 'Personal Identity and Memory', Journal of Philosophy, 56 (1959) pp.868-82.
47. Shoemaker, S. Self-knowledge and Self-Identity (Cornell University Press, New York, 1963).
48. Shoemaker, S., 'Self-Reference and Self-Awareness', Journal of Philosophy, 65 (1968) pp.555-67.
49. Shoemaker, S., 'Wiggin on Identity', Philosophical Review, 78 (1970) pp.529-44.
50. Smart, B., 'How can Persons be Ascribed M-Predicates?' Mind, 86 (1977) pp.49-66.
51. Smart, J.J.C., 'Sensations and brain processes', Philosophical Review, LXVIII (1959).
52. Strawson, P.F., Individuals (Methuen, London, 1959).
53. Strawson, P.F., 'Critical Notice : Philosophical Investigations' Mind, LXIII (1954) pp.70-99.
54. Vesey, G. Personal Identity (Macmillan, London, 1974).
55. Wiggins, D., Identity and Spatio-temporal Continuity (Basil Blackwell, Oxford, 1967).
56. Williams, B., 'Bodily Continuity and Personal Identity - a Reply', Analysis, 21(1961) pp.43-8.
57. Willaims, B., 'Mr Strawson on Individuals', Philosophy, 36 (1961) pp.309-332.
58. Williams, B., 'Personal Identity and Individuation', Proceedings Aristotelian Society, 57 (1956-7) pp.229-52.

59. Williams, B., Problems of the Self (Cambridge University Press, Cambridge, 1973).
60. Williams, B., 'The Self and the Future', Philosophical Review, 79(1970) pp.161-80.
61. Wittgenstein, L., Philosophical Investigations (Basil Blackwell, Oxford, 1953).
62. Wittgenstein, L., The Blue and Brown Books (Basil Blackwell, Oxford, 1958).