Complex Persons: A Holistic Solution to Personal Identity

Doctoral Thesis: Philosophy

University of Tasmania

Pauline Enright 2002

Table of Contents

TABLE OF CONTENTS **0**

CHAPTER 1 INTRODUCTION **3**

1.1 THESIS OUTLINE 3
1.3 PERSONAL IDENTITY THEORIES 9
1.4 RELATION R 13
1.6 PERSONHOOD 26
1.6.1 Summary of Personhood 33
1.7 HOLISTIC PERSONS 37
1.8 COMPLEX PERSONS 38

CHAPTER 2 UNCOMMON PERSONS **43**

2.1 INTRODUCTION 43
2.2 THOUGHT EXPERIMENTS CONSIDERED 45
2.3 THOUGHT EXPERIMENTS ASSESSED 48
2.3.1 Summary of Assessment 57
2.4 LOCATING THE DIFFICULTIES 58
2.5 SUMMARY 64

CHAPTER 3 EXTERNALISED MINDS **66**

3.1 INTRODUCTION 66
3.2 THE PROBLEM WITH INTERNALISM 68
3.3 THE CASE FOR EXTERNALISM 72
3.4 PERCEPTION 76
3.5 PERCEPTION OF THE ENVIRONMENT 77
3.6 PERCEPTION AND AFFORDANCE 88
3.7 VISUAL PERCEPTION 93
3.8 SUMMARY 106

CHAPTER 4 PSYCHOLOGICAL COHERENCE **109**

4.1 INTRODUCTION 109
4.2 QUASI-MEMORY 113
4.3 PSYCHOLOGICAL ATOMISM 121
4.4 PSYCHOLOGICAL HOLISM 128
4.5 MEMORY THEORIES 134
4.6 MEMORY RESEARCH 138
4.6.1 Memory Taxonomies 139
4.6.2 Reconstruction 142
4.6.3 Eye-witness Testimony 143
4.6.4 Flashbulb Memories 145
4.6.5 Summary Of Memory Research 146
4.7 FALSE MEMORIES 147
4.7.1 Memory Illusions 148
4.7.2 False Recognition 154
4.8 SUMMARY 160

CHAPTER 5 INDISCRETE BODIES **162**

5.1 INTRODUCTION 162
5.2 IRRELEVANT BODIES 169
5.3 DISEMBODIED MEMORIES 170
5.4 THE UNITY OF CONSCIOUSNESS 176
5.5 BRAIN-BODY INTEGRITY 182
5.6 SUMMARY 190

CHAPTER 6 THE SELF AS DYNAMIC UNITY 193

6.1 INTRODUCTION 193
6.2 THE 'NO-SELF' VIEW 197
6.3 SEARCHING FOR THE SELF 200
6.4 COGNITIVE UNITY 203
6.5 EMBODIED UNITY 210
6.6 THE IMPORTANCE OF MOVEMENT 215
6.7 THE IMPORTANCE OF INTEGRATION 220
6.7.1 Habituation 224
6.7.2 Olfactory Discrimination 225
6.7.3 Phantom Limbs 227
6.7.4 Imitation 230
6.8 REMEMBERING ONESELF 232
6.8.1 The Beginnings of Self-Identity 233
6.8.2 The Emergence of the Cognitive Self 235
6.8.3 Autobiographical Memories 238
6.8.4 Narrative Identity 240
6.9 SUMMARY 246

CHAPTER 7 CONCLUSION 248

7.1 CENTRAL CLAIMS 248
7.2 FUTURE DIRECTIONS 251

BIBLIOGRAPHY 253

Chapter 1 Introduction

1.1 Thesis Outline

The problem of personal identity is the problem of specifying what makes a person at one time, the same person at a later time.[1] For example, what is it that makes me the same person that I was twenty years ago, or that I will be in twenty years' time? According to one current view, personal identity over time is preserved in 'overlapping chains' of psychological connectedness.[2] These overlapping chains contain memories, beliefs, and other mental items. Personal identity accounts are accounts of psychological chains and the items within these chains. These accounts require reference neither to the items' ownership, nor to other items in the same chain. Known as the 'psychological approach,' this view also holds that mental items do not depend crucially for their identity on other features, such as particular bodies or brains, or on underlying entities such as souls or substantial selves. While the accumulation of mental items depends on experiencing subjects, the subjective apprehension of experience requires no special or substantial underpinnings. Theories embracing the psychological approach are criterial, as they specify the relevant criteria under which personal identity is preserved, and are known as versions of the *psychological continuity criterion.* Thought experiments, in which mind parts become disconnected from bodies or brains, or in which persons are 'reduplicated,' are frequently used to support the psychological continuity criterion, as they claim to show that personal identity is preserved in the mind, to the exclusion of the body or other external factors. In this thesis, I challenge the psychological continuity criterial approach, on the grounds that its portrayal of personal identity is distorted and misleading.

---

[1]The problem of personal identity is more fully developed throughout this chapter, especially in Section 1.2.

[2]For example, see (Parfit 1984), p 206.

My challenge addresses five key problems in the psychological continuity criterion that are referred to in the above account, but which are set out more precisely below:

- The excessive reliance on oversimplified thought experiments to analyse personal identity.
- The view that mind contents are adequately described by just looking inside the mind.
- The view that minds are atomistically structured, and that mind contents are not individuated by ownership.
- The view that the body is incidental to personal identity.
- The view that the selves of experiencing subjects are no more than mere collections of thoughts and experiences.

In addressing these five problems, I present some ideas which I see as crucial to personal identity, and which should be incorporated in any account of personal identity if such an account is to be accurate and comprehensive. It is my view that contrary to the criterial approach, personal identity must recognise that persons are complex and dynamic entities, in which their many components operate in an integrated manner, thereby producing a holistic structure. Unlike the criterial view of persons, this view requires that a person is understood as a whole, and not merely as a sum of theoretically separable parts. A comprehensive account of personal identity, therefore, is one which recognises this holistic conception of persons, and accounts for them accordingly. To support the holistic conception of persons, the central claims of this thesis are as follows:

- That due to its reductionist, criterial approach, the psychological continuity criterion is inadequate to account for personal identity.
- That a sound approach to personal identity must respect the complex, dynamic, holistic, non-reductive nature of persons.

Supporting these two main claims are five pairs of supporting claims, which amplify the five problems outlined above:

- Thought experiments are inadequate to comprehensively analyse personal identity. In addition to such philosophical analysis, empirical research concerning actual persons in real situations is required.

- An internalist view of mental content is inadequate to account for the mind. Minds are formed and operate only in virtue of features external to the mind, and consequently, the constitution of minds cannot be understood independently of the environments in which they are located.

- An atomistic, 'impersonal' characterisation of mental content misrepresents the mind's structure and mode of operation. Minds are predominantly holistic in structure and functioning, and mental states are individuated by ownership.

- The neglect of the body by the psychological continuity criterion produces misleading conclusions about personal identity. The body plays a decisive role in psychological continuity, and therefore, also in personal identity.

- The reduction of the self to a mere sum of experiences leaves subjectivity and experience inadequately explained. The self is a dynamic unity, which is essential to subjectivity and experience, and consequently, also to personal identity.

These five supporting claims will be argued for throughout the next five chapters. While my arguments apply to the psychological continuity criterion in general, most of the discussion will focus on the version associated with Derek Parfit. Parfit is prominent among theorists who argue in favour of psychological continuity. Bodily continuity is also argued for by some current theorists, such as Peter Unger and Eric Olson.[3] Due to the complex nature of their theories, their work cannot be addressed in detail here. Although psychological continuity theories are the main focus of this thesis, a point that emerges is that, due the integrity of the mind/ body relation, the separation of the mind and the body is as illegitimate for bodily continuity theorists as it is for psychological continuity theorists.
Indeed, as I will show, any theory which assumes that the mind and the body operate autonomously, without regard for their integrity, or for their locatedness within a specific place or environment, does not take seriously the issues which personal identity actually involves.

[3]See (Olson, 1997) and (Unger, 1990).

Parfit's work is chosen as the main focus of this thesis as it has received much criticism,[4] and because it is more fully developed than most other versions, although many of its key ideas would most likely apply to these other versions were they to be equally developed. A feature of Parfit's version is his detailing of the ethical ramifications which would follow were the psychological continuity criterion to be adopted. I find these ethical stances questionable. Although Parfit claims that more concern for others and more personal liberation would follow, my view is that the opposite would occur. I suspect that with the neglect of crucial features involved in personal identity, such as the environment and the body, and with the reduced conception of the self, concern for others would decrease, responsibility would decline, and personal autonomy and self-esteem would diminish.

The remainder of this chapter will proceed by first explicating more clearly the 'problem' of personal identity. Some responses to this problem, including major aspects of Parfit's theory, will then be considered. Several difficulties which arise from Parfit's theory will then be discussed. I will then consider the issue of 'personhood,' and outline the holistic approach to personal identity for which I shall argue throughout this thesis. Finally, the issues which occupy the remaining chapters will be summarised.

1.2 The Problem of Personal Identity

Personal identity involves the set of conditions under which a person remains the same person over time, that is, the same person at a later time as at an earlier time.[5] Under normal conditions, being the same person over time involves being physically and psychologically continuous with an earlier person, that is,

---

[4]See, for example: (Brueckner 1993), (Cartwright 1993), (Curzer 1991), (Lewis 1976), and the many articles referenced throughout this thesis.

[5]The material in this section is based on reading from a variety of sources, including: (Baillie 1990); (Lyon 1988); (Maddell 1981); (Noonan 1989); (Nooonan 1993); (Parfit 1984), and (Shoemaker 1984).

maintaining continuity of the same body and mind. But closer consideration reveals that bodies and minds are complex, and difficult to analyse. Bodies have parts, such as cells, limbs, organs, and brains. Minds involve 'components,' such as thoughts, beliefs, emotions, and memories. Are all these parts essential to the preservation of personal identity, or are some essential while others are not? Some personal identity theorists hold that problem cases arise where body or mind parts are missing, or where bodies and minds could become separated. They claim that in such cases, we need to know what the conditions are under which personal identity is maintained. The problem of personal identity is just the difficulty of stipulating what those conditions are. Harold Noonan describes the problem as:

> the problem of giving an account of the logically necessary and sufficient conditions for a person identified at one time being the same person as a person identified at another. Otherwise put, it is the problem of giving an account of what personal identity over time necessarily consists in, or as many philosophers phrase it, the problem of specifying the criterion of personal identity over time (Noonan 1989), p 2.

The problem of personal identity is a problem of persistence through change. It concerns the question of what permits persons to remain the same persons over time, when every physical cell and every mental state is constantly changing. The notion of 'identity' requires explanation. 'Identity' can refer to what is often understood as' type' identity or 'token' identity. Type identity refers to the identification of objects according to their possession of certain types or groups of characteristics. For example, one could say that what identifies Socrates is that he is a 'rational animal.' Similarly, one could say that what identifies Aristotle is also that he is a 'rational animal.' In both cases, Socrates and Aristotle are the same *type* of thing, and so have the same *type* identity. Socrates and Aristotle, however, do not have the same *token* identity. They are each different instances of type identity, and thus, each have a different token or individual identity. So,

although they are identical to each other in some respects, they are also *numerically* distinct from each other. Personal identity concerns token or individual identity, as it is concerned with the identities of particular, individual persons.

Token identity is, then, the identity of one thing over time and at a time, according to which the thing in question is distinct even from those other things which are qualitatively the same, and in which the thing concerned occupies a different spatio-temporal location from those other things. Thus we can distinguish numerical from qualitative identity, as it is numerical identity that is normally at issue in the discussion of personal identity. The identity of particular objects, however, is not always clear-cut, as there can be problem cases, in which identity is difficult to pin down or in which identity appears to change. For example, the 'contents' of a river are always moving, yet we understand the river to always be the same river. A piece of clay appears to change its identity when it becomes a statue, yet the material of both objects is the same. Similarly, a signpost appears to change identity when its function changes and it becomes a see-saw, yet the object itself in both cases is exactly the same object. It can thus be seen that identity is not always simple and easy to determine.

Living objects also change, as their interaction with the world generates change and decay in the particles out of which such objects are constituted. Where the change is normal, non-human living objects are taken to maintain identity over time. Persons are more complex, as they comprise both biological and psychological lives. Because both physical and mental lives are subject to constant change, the locus of personal identity is obscured and problematic. This problem is now explored in more detail.

1.3 Personal Identity Theories

Personal identity theories demonstrate a variety of responses to the problem outlined above. These theories are traditionally held to have originated in the work of John Locke, although the mind-body dichotomy on which they are typically based goes back to Plato. Locke holds that persons are primarily psychological entities, whose identity is preserved in rationality and consciousness:

> 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 thoughts, so far reaches the identity of that person; it is the same self now it was then; and it is by the same self with this present one that now reflects on it, that that action was done (Locke 1959), 2.27.11.

According to Locke's view, personal identity over time is realised in virtue of continuing successive conscious states, which are usually understood as memories, beliefs and other similar mental items. Locke's theory is foundational to many contemporary personal identity theories. These can be divided into *non-reductionist* and *reductionist* theories.

Non-reductionist theories hold that personal identity cannot be fully accounted for. Typically (though not necessarily),[6] non-reductionist theories are dualistic in nature, and are explicitly based on Cartesian ontology, where mind and body are regarded as fundamentally separate entities. A person is regarded as a unique entity which is supplementary to the body and brain.[7] The identity of a person is not reducible to a specifiable set of facts, but is based on either a Cartesian soul,

---

[6]As will become apparent, I offer in this thesis a non-reductionist view of personal identity which is *not* based on Cartesian ontology.

[7]Among current theorists who hold non-reductionist views are Geoffrey Madell, Roderick Chisholm, and Richard Swinburne. See, for example: (Maddell 1981); (Chisholm 1969); and (Swinburne 1984) .

or on a further, unanalysable or undiscoverable fact.[8] Thus, a complete description of a person cannot be given.

Personal identity is always determinate and specific, as opposed to the identity of other living objects, which is sometimes 'vague.'[9] For example, the identity of a plant may seem 'vague' due the constant change in the material out of which the plant is composed. The constant change of matter also occurs in humans, but on the non-reductionist view, the personal identity of humans is taken to be unaffected by such change, and to remain strict and unaltered. Personal identity is what matters to a person's survival, and to a person's ethical commitments. Two main reasons for holding a non-reductionist view are the commitment, frequently on religious grounds, to an immaterial (often immortal) soul or substance, and the disillusionment with the paradoxes generated in reductionist accounts, such as the problem of possible person-reduplication.[10]

---

[8]This view is traditionally attributed to Butler - see (Maddell 1981), p 107; (Noonan 1989), p 64; (Parfit 1984), p 222; and (Shoemaker 1984), p 75. Butler claimed that analysis of personal identity reached an end-point beyond which it could not go: 'But, though we are thus certain that we are the same agents, living beings, or substances, now, which we were as far back as our remembrance reaches; yet it is asked, whether we may not possibly be deceived in it? And this question may be asked at the end of any demonstration whatever, because it is a question concerning the truth of perception by memory.

And he who can doubt, whether perception by memory can in this case be depended upon, may doubt also, whether perception by deduction and reasoning, which also includes memory, or indeed whether intuitive perception, can. Here, then, we can go no farther. For it is ridiculous to attempt to prove the truth of those perceptions, whose truth we can no otherwise prove than by other perceptions of exactly the same kind with them, and which there is just the same ground to suspect; or to attempt to prove the truth of our faculties, which can not otherwise be proved than by the use of means of those very suspected faculties themselves' (Butler 1867), p 198.

[9]The origin of this view is attributed to Butler. He referred to the non-reductive view of personal identity as 'strict and philosophical', as opposed to views about the identity of other objects, which were 'loose and popular'. He found Locke's treatment of personal identity 'hasty', and was at pains to draw a distinction between the identity of persons, and that of other objects. To this end he contrasted the identity of plants to that of humans: 'And therefore, when we say the identity, or sameness, of a plant consists in a continuation of the same life, communicated under the same organization, to a number of particles of matter, whether the same or not, the word *same,* when applied to life and to organization, cannot possibly be understood to signify what it signifies in this very sentence, when applied to matter. In a loose and popular sense, then, the life, and the organization, and the plant, are justly said to be the same, notwithstanding the perpetual change of the parts. But in a strict and philosophical manner of speech, no man, no being, no mode of being, no any thing, can be the same with that with which it hath indeed nothing the same. Now, sameness is used in this latter sense when applied to persons. The identity of these, therefore, cannot subsist with diversity of substance' (Butler 1867), pp 194-195. These views of Butler are seminal to some of the arguments raised in the personal identity debate.

[10]This problem is addressed in Chapter 2.

Reductionist theories hold that personal identity is reducible to a set of specifiable facts. Describing these facts is said to produce a full description of a person. These facts concern either bodies or brains, or the states of bodies or brains. Materialist conceptions of the mind traditionally lend themselves to reductionist accounts of personal identity, as these accounts do not attempt to ground personal identity in unknown souls or non-material entities. Reductionist theories provide criterial accounts which stipulate the necessary and sufficient conditions required for a person at one time to remain the numerically same person at another time. Criteria stipulated for personal identity *over* time (diachronic identity) are thus derivable from, and indeed, necessarily limited to those stipulated for personal identity *at* a time (synchronic identity). Criterial accounts are of two types, and refer to either *physical* or *psychological* continuity.

Physical continuity theories hold that it is the continuity of all or part of the human body which accounts for the identity of persons over time. The degree to which psychological states are also implicated varies according to the particular theory. There are three main types of physical theories. *The Bodily Criterion* stipulates that a person is the same person at one time as she is at another time if and only if she has the same body. Minute cell change of living tissue is permitted, providing that the change is normal, and occurs gradually and naturally. *The Brain Criterion* holds that personal identity resides in the brain, rather than in the whole body. If a person's brain was transferred to a different body, identity goes with the brain, rather than with the body. Psychology is important, but only because it resides in the brain. *The Physical Criterion* holds that personal identity is retained in amounts of physical matter less than that of a complete human brain. The amount of matter ranges from slightly less than a whole brain, such as when a small amount of brain is removed, to minuscule amounts, such as a single cell. Brain-split operations, both real and imaginary are appealed to in support of this

view, as these operations have convinced some theorists that incomplete brains could maintain the psychology of a whole person.[11]

Psychological continuity theories hold that it is the continuity of a person's psychological states which maintains personal identity over time. Person *A* at t_1 is the same person as person *B* at t_2 if and only if the psychological states of person *A* are appropriately connected to the psychological states of person *B*. Appropriate connectedness can refer to 'any causal links between past factors and present psychological states' (Noonan 1989), p 13.

While psychological states are instantiated in some body or other, the body itself contributes nothing essential, either to the states or to personal identity. Psychological continuity theories specify criteria under which personal identity is preserved, and are known as versions of the *psychological continuity criterion*. The psychological continuity criterial approach generates certain problems. In particular, this approach is vulnerable to producing duplicate accounts of personal identity. It is in principle possible that the *same* description could be given to more than one set of mental states, theoretically resulting in the apparent conflation of several persons into one person. This problem is exacerbated by the consideration of imaginary cases in which persons are 'duplicated,' or in which brains are split and put into different bodies.

The attempt to overcome these and similar problems has resulted in several versions of the theory being produced. Each version attempts to overcome the duplication problem in a different way. A major problem with these responses is that they are incapable of being subjected to empirical scrutiny. It is simply

[11]This issue is developed in Chapter 5. There are many descriptions of brain operations in personal identity literature. See, for example (Nagel 1979), and (Unger 1990), or for more seminal material, (Gazzaniga 1970), and (Gazzaniga 1978).

assumed that they can be accepted at face value.[12] Parfit's is one such version and is now considered.[13]

1.4 Relation R

To distinguish his version of the psychological continuity criterion from other reductionist theories, Parfit refers to it as 'Reductionist,' with a capital 'R.'[14] His Reductionist view holds that personal identity is reducible to psychological continuity, and comprises no underlying entities such as souls, Cartesian Egos, or substantial selves. Much of Parfit's reasoning is based on thought experiments, which involve brain or part-brain transfers, or the supposed artificial duplication of persons. When these transfers or duplications occur, the body which is discarded is taken to have no significance to the maintenance of the personal identity concerned. Parfit's thought experiments underpin his conclusion that personal identity holds in virtue of non-branching psychological continuity, defined as 'overlapping chains:'

> *Psychological connectedness* is the holding of particular direct psychological connections.
> *Psychological continuity* is the holding of overlapping chains of *strong* connectedness (Parfit 1984), p 206.

Psychological continuity is defined in terms of 'overlapping chains' because on their own, 'direct psychological connections' are not transitive relations, and thus would be unable to maintain identity. Parfit claims that while he is strongly

---

[12]See, for example: (Lewis 1976), in which Lewis argues for the *'Multiple Occupancy thesis.* ' This thesis holds that in the cases of supposed fission (person reduplication), that new persons are not really created, but have existed all along, and have only become spatio-temporally evident after the supposed fission procedures have occurred (Noonan 1993), p xvii, and also (Nozick 1981), especially pp 29-114, in which Nozick argues for the *'Closer Continuer Theory.'* This theory holds that the person which survives is the person who is closest to, or most like, the original person. Persons other than the 'closest continuer' who are supposed to have been produced are said to not maintain identity with the original person.

[13]Parfit's theory is outlined only, as the relevant areas are expanded in future chapters.

[14]Thus, when referring to 'Reductionism' as opposed to 'reductionism,' I will be specifically referring to Parfit's theory.

connected to himself yesterday, and he was then strongly connected to himself the day before, and so on, it does not follow that he is strongly connected to himself as he was twenty years ago. He claims that there could be very few connections for most people over long periods of time. 'Overlapping chains,' on the other hand, are transitive, and permit identity to persist over extended periods of time.

But even so, overlapping chains require sufficient connections between earlier and later times for identity to persist. Parfit claims that ‘strong connectedness’ is the retention of at least half the number of direct connections which hold in a single day. Identity is strongest when a large number of connections are held, and is weaker as the number lessens. Where less than half are held, identity fails to be realised (Parfit 1984), p 206.

For Parfit, causes must be present between the elements which comprise psychological continuity. This means that the components in a person's psychological chains must have causal connections between them of some kind. Causes are important, as it is the causal connections between items which permit identity to be retained. Three types of cause are possible, *Narrow*, *Wide,* or *Widest.* 'Narrow' cause is normal cause, such as the retention of one's memories in the normal way, or the normal maintenance of a continuous spatio-temporal path through the world by one's body. 'Wide' cause is any reliable cause, such as a partial or complete brain-transfer. 'Widest' cause is any cause whatsoever, such as the complete reduplication of a person. Parfit claims that the 'Widest' criterion is best, as, even if it is not 'true continuity,' it is just as good. We should accept the 'Widest' criterion because there could be instances where we would accept non-normal causation in relation to bodily functions, and we should, therefore, accept non-normal causation in relation to mental functions. For example, if a normally-blind person were fitted with an artificial device which enabled a form of vision to

occur, we would accept that this was 'as good' as normal vision, even though it was not caused in the normal way:[15]

> I shall argue that the two Wide Psychological Criteria are both better than the Narrow Criterion. A partial analogy may suggest why. Some people go blind because of damage to their eyes. Scientists are now developing artificial eyes. These involve a glass or plastic lens, and a micro-computer which sends through the optic nerve electrical patterns like those that are sent through this nerve by a natural eye. When such artificial eyes are more advanced, they might give to someone who has gone blind visual experiences just like those that he used to have. What he seems to see would correspond to what is in fact before him. And his visual experiences would be causally dependent, in this new but reliable way, on the light-waves coming from the objects that are before him. Would this person be *seeing* these objects? If we insist that seeing must involve the normal cause, we would answer No. But even if this person cannot see, what he has is *just as good as* seeing, both as a way of knowing what is within sight, and as a source of visual pleasure. If we accept the Psychological Criterion, we would make a similar claim. If psychological continuity does not have its normal cause, some may claim that it is *not* true psychological continuity. We can claim that, even if this is so, this kind of continuity is *just as good as* ordinary continuity (Parfit 1984), pp 208-209.

Parfit argues that because we should accept that non-normal causation of psychological continuity is just as good as normal causation, we should also accept that even a fully duplicated 'person' would be as good as an original person. Providing that only a single 'person' resulted from the duplication (or brain-transfer) process, psychological continuity with the original person would be maintained. Parfit designates the relation involved in this type of psychological continuity as *Relation R,* and defines this relation as 'non-branching psychological continuity with any cause' (Parfit 1984), pp 199-217.

For Parfit , the commitment to Relation R holds certain metaphysical consequences for the status of persons. Because personal identity is

[15]As Parfit defends his claims for causal continuity with *any cause* with this argument, his defence is presented in full.

characterised in terms of a certain amount of overlapping chains, the maintenance of personal identity is a matter of having sufficient of these chains. When the amount of chains a person owns drops below a certain level, that person ceases to exist, and consequently, the person that subsequently exists is a different person to the one who existed earlier. Thus, an aged person whose mental constituents have undergone substantial change since her youth will be a different person to the person she was many years earlier. On this view, the life of a single body could embrace the existence of more than one person at different times in its life-history. Parfit holds that the difference between these persons at different times could be as great as the difference between different-bodied persons at the same time. In some instances, it could not be known when a person changed from being one person to being a different person. Parfit holds that this means that personal identity is often indeterminate, and is, therefore, less important that we might normally think. It is largely insignificant whether an individual is one person, or several different persons over time. Psychological continuity is more important than personal identity.

Parfit also holds that these metaphysical consequences have certain ethical ramifications. Different persons who are connected to each other within the lifetime of a single body may not be accountable for each other's actions. These persons exist at different times, and are 'series persons' (Parfit 1984), p 298. Series persons are not responsible for each other's actions, commitments, promises, sufferings, crimes, and so on. Moral responsibility between series persons diminishes in proportion to reduced psychological connectedness. Parfit gives some examples of ways in which persons' lives could be affected by this changed conception of moral responsibility.

First, a wife could make a promise to her husband when he was a young man, who had a certain set of values. Later, the wife could revoke that promise, as she

might consider that her husband's values had changed, and that he had therefore become a different person to the one to whom she made the promise. She could hold that the young man to whom she has loyalty no longer exists, and that her now middle-aged husband is, therefore, owed no such loyalty (Parfit 1984), pp 327-329.

Second, it may not be possible to compensate suffering which occurs early in life by gains received later in life. A person who suffers at one time may no longer exist by the time the gains are received:

> I argued earlier that if we move from the Non-Reductionist view to the Reductionist view, it becomes more plausible to claim that there is less scope for compensation within the same life. Thus it is more plausible to claim that great burdens imposed upon a child cannot be compensated, or fully compensated, by somewhat greater benefits in this child's adult life (Parfit, 1984), p 346.

The life of these persons does not coincide with the life of a single body, consequently, the life-time of a single body cannot be conceptualised as a complete whole, as if it was the life-time of only one person.

Third, a person who commits crimes at an earlier time may not be the same person who receives punishment for those crimes at a later time. Parfit argues that if a criminal is 'less connected' to himself than he was at the time of the crime, he deserves less punishment. When connections are very weak, maybe no punishment at all is warranted (Parfit 1984), pp 326-346.

Fourth, because series persons could be as separate from each other as spatially discrete persons, distributive justice between persons should be altered. Parfit claims that principles of distributive justice should be accorded 'more scope and

less weight.' He claims that it makes sense to be more concerned with justice across a community, than with justice for individual persons. Issues of moral responsibility, compensation or commitment, should be viewed similarly between different bodied persons, as they are between successive persons in a single body. Because the boundaries between different bodied persons and different series persons are not distinct, it would be irrational for me to be especially concerned for my own future, rather than for someone else's future.

Self-interest theories should be replaced by more impersonal theories, such as the 'Revised Self-Interest Theory,' as Parfit terms it. According to this theory, it may not be irrational to do things against one's own interests. For example, it would not matter if a person suffered hardship as long as *someone* benefited. It does not matter if the person who suffered is not the person who gains. On the Reductionist view, identity is less deep than on non-reductionist views, so we should be more concerned with the quality of experiences, than with whose experiences they are. Parfit claims that we ought to be Reductionists, and hold to the above moral beliefs. He claims that these beliefs can generate more concern for others, and can be personally liberating (Parfit 1984), pp 346-347.

1.5 Problematic Persons

In providing Relation R, Parfit has responded to the psychological continuity criterion's inability to provide uniquely determining personhood characterisations or personal identity descriptions. But the consequence of Relation R is that the status of personhood is diminished to the extent that personhood itself is secondary to the characteristics of which it is composed. More specifically, because the psychological continuity criterion is incapable of capturing the precise elements of personal identity, the psychological states which are deemed to characterise personal identity have become more important than personal identity

itself. As Parfit states, significant metaphysical and ethical ramifications follow from this stance. Both of these are seriously problematic.

Parfit claims that the commitment to Relation R should lead to an improvement in moral consciousness. He argues that because the boundaries between different persons are blurred, the concern we have for ourselves would lessen, and the concern we would have for others would increase. But it is not clear that this would necessarily be the case. The concept of series persons means that responsibility between earlier persons and later persons is diminished, and in some instances, is lost altogether. This view of responsibility invites the devaluing of the very concept of responsibility. If persons thought of themselves as series persons and knew that they may not in the future be held responsible for their actions as they would no longer exist, they may care less about what those actions are than they would do otherwise.

Such a concept of personhood could encourage capricious or arbitrary promise-making and promise-breaking. It could also generate a careless approach to one's actions and one's attitudes towards others. Often the way we behave and choose our actions is associated with what we perceive to be the interests of ourselves and of other persons who are affected by such actions. But the idea that we or our friends and families may no longer exist when such interests are realised is likely to make us care less about what those interests are. Why should I care if I am mean and unjust to others now, if later I can later tell myself it was not me who was once mean and unjust, but someone else? The notion that I would care for *someone* in the way that I care for myself is simply not feasible, as it is only in virtue of caring for myself that my ability to care for others is possible.[16]

[16]Imagine that you never had any needs or concerns. How would you know what it was like to want or need something, or to care for someone if you had never experienced the meaning of 'care'? How would you understand the needs of others? How could you care for them, or put them before yourself, if you had no conception of need, or of meeting your own needs?

Another problem with series persons concerns the elderly and sufferers of diseases such as Alzheimer's disease. Due to the loss of psychological connections, these persons would be considered to be no longer be the same persons they once were, and in some instances, they would be considered to no longer be persons at all. One's elderly relatives may not be the same persons they were when one was a child. At a time when your mother may most need your care and attention, she may not be the same person who was once your mother. Many people would find this notion of identity disturbing.

Parfit might say that it would not matter if the person was not your mother, as long as you cared for someone. But this view is not realistic, and takes no account of our intuitions about our family and other close relations. In most instances, we instinctively conceive of our family (or surrogate family) relations as being unique, and as being, in an important sense, unaffected by change. We do not regard our relatives as being different persons just because they lose their memories, or because they change their ideas over the years.

Even if we do not get on well with our families, we take for granted that they are, in some sense, an aspect of our lives, which, in normal circumstances, could not be arbitrarily discarded or exchanged. Family relations are also unique in that they are the basis of our relations with others. Being the primary form of our relations, they are *foundational* to our relations with others, and thus cannot in some arbitrary way become secondary to that foundation. Families are also part of the structure of society, and an important basis of social relations. If family relations were subject to the sort of breakdown envisaged by the psychological continuity criterion, social relations, and with them societies, would be vulnerable to breakdown also.

Furthermore, the diminishing or removal of responsibility and punishment for crimes, on the basis of a reduced psychology, could lead to widespread injustice. Because the duration of series persons depends solely on the amount of psychological change which occurs, it effectively means that whether or not someone at a later time is held responsible for a crime committed at an earlier time depends solely on the *quantity* of change involved in that person's psychology, rather than on the *quality* of change involved. This seems grossly unjust. It means that if persons (as traditionally understood) were at one time involved in horrific war crimes, they could, at a future time be excused the responsibility for those crimes if their psychology had sufficiently changed, *regardless of whether the change which had occurred involved remorse, character improvement, or moral growth, or victims had been compensated.*[17]

Similarly, a person who commits corporate fraud at an earlier time and lives off the fruits of his profits on a luxurious island at a later time, given the right amount of 'memory loss,' could fail to be held accountable for his former actions, *regardless of whether he felt remorse or had compensated his victims.*

The above attitudes towards moral responsibility go against our intuitions, and seem neither ethical nor just. They make neither for an ethical individual, nor for an ethical society. If the principles generated by series persons were universalised, society as a whole would suffer. Without realistic and feasible distinctions which permit persons to be held accountable for their actions, personal autonomy would diminish (if not lose) significance, and the impetus for social and moral reform would be inhibited. If it does not matter *who* commits particular actions, it can hardly matter *who* is affected by them, and so it cannot matter *what* actions are committed. These ethical stances are controversial, and

[17]It is not enough to say: Well, if someone had changed that much, of course they would be reformed.' It is not at all certain that this would be the case. All that is required is the loss of a certain amount of memories, not a character change or compensatory actions. It is the principle itself - quantity v quality - that is at issue.

have the potential to affect society considerably were they to be adopted as general principles. As I believe that these effects would be negative, the view of persons which leads to such stances should be questioned. In my view, these unsatisfactory ethical stances arise from the conception of personal identity associated with Relation R.

The major fault of this conception is that it neglects to acknowledge many aspects of personal identity, which, although not explicitly specified by the psychological continuity criterion, are nevertheless implicated by it. In other words, conceiving of personal identity in terms of psychological states alone, without taking account of what is involved in those states, leaves many important aspects of personal identity unaccounted for. This neglect ultimately leads to the loss of personal identity and to the incurring of certain ethical ramifications which, in my opinion, are totally unacceptable. It is the gap in analysis that concerns the present thesis, as, in my view, this gap is responsible for an account of personal identity which is inadequate and misleading. Of the areas of neglect involved, five important problems in Parfit's account of personal identity were referred to earlier, and are now outlined in more detail.

First, Parfit places undue emphasis on thought experiments to elucidate personal identity. Like those of other personal identity theorists, Parfit's thought experiments contain fictitious scenarios, in which extraordinary things happen, and in which extra-ordinary persons are produced. Concerns about the identities of these extra-ordinary persons are then used to inform about the identities of normal persons. It is not clear, however, that the identities of extra-ordinary persons have much in common with the identities of normal persons, as the scenarios which produce extra-ordinary persons give little exploration or time to the concerns of real persons in real life. This lack of concern with real persons

results in a distorted and inappropriate assessment of what personal identity involves.

Second, Parfit fails to recognise the significance of causal factors. The psychological continuity criterion stipulates that the contents of mental chains are essential to personal identity, but does not care how these contents are caused. The *Widest Criterion* holds that 'any cause will do,' which means that the *actual* cause may be entirely ignored. This means that factors *outside* the mental chains which contribute to the chains' contents are effectively excluded from the account of mental content. This is misleading, as what goes on 'inside' the mind is crucially affected by what goes on 'outside' the mind.[18] Parfit is little interested in external factors, and concentrates on the mental chains themselves and what happens to them, rather than on where they came from.

Third, in giving an account of personal identity in terms of overlapping chains, Parfit characterises minds and mental items atomistically. Personal identity is maintained according to the number of connections between mental items. Neither the ownership of these items, nor the way these items are connected to each other is taken to be of importance. But this view of the mind's structure is inaccurate, as it misconstrues the nature of the mind and its mode of operation. This view misses the point that the ownership of mental items and the relation between different mental items are essential features of the mental items themselves. To be accurate, an account of personal identity must recognise these essential features, and must include them when accounting for minds and the contents of minds.

[18] I place 'inside' and 'outside' in inverted commas because I think, as will be shown, that this distinction is artificial and misleading. As I shall argue, especially in Chapter 3, what is 'inside' our minds is ultimately what is 'outside' them also.

Fourth, when accounting for mental states, Parfit neglects the body's significance. Whether a person has one body or another body is not taken to be of interest to the mental life of that person. Different persons can occupy the same body at different times, without their identity being compromised. This means that any influence which the body has on mental life goes unrecognised. The kind of bodies persons have is taken to have no bearing on the kind of minds they have, nor on the kind of thoughts they think. This is not right, as the particular body a person has is crucially related to that person's mind. We cannot take for granted that minds would be the same, were they to be 'placed' in different bodies.

Fifth, the Reductionist stance on the self amounts to an unacceptable trivialisation of the self. Reductionism conceives the self to be no more than a formal requirement of experience, but does not expand on what this formal requirement involves or implies. Parfit maintains that experience requires no underlying entities, such as unknown souls or Cartesian Egos, and that, therefore, we should accept Reductionism and reject such types of entities. This view seems to conflate two issues: the question of the existence of an underlying entity beyond the reach of empirical research, and the question of the existence of a detectable empirical self, which, although substantially constituted by experience, nevertheless, is more than a mere sum of experiences. The confusion arises from viewing experience from a third-person perspective only. When the first-person viewpoint is taken into account, it becomes evident that a subject's first-person knowledge of herself is a crucial condition of experience. The relevance of the first-person viewpoint to the preconditions of experience is not taken seriously in the above account, and this leads to a misleading conception of the self.

The above deficiencies in Parfit's account demonstrate the psychological continuity criterion's failure to adequately account for personal identity. This failure is exemplified in the fact that by trying to account for personal identity in terms of

psychological states alone, the psychological continuity criterion loses the persons it started with, and as a result, undermines the concept of moral responsibility between different persons, and possibly across society as a whole. As this view of personal identity and its consequences is controversial, the analysis which leads to such a view, namely *criterial* analysis, warrants investigation.

Criterial analysis of personal identity operates by dissecting persons, looking at their parts, ascertaining what the parts are used for, and then determining which parts are essential to personal identity. Accounts of personal identity formed on this basis then stipulate the relevant parts which encapsulate personal identity, to the exclusion of other parts which are seen as unessential, and as theoretically expendable or replaceable. A major problem with this type of analysis, however, is that in conceptualisinng the parts which are candidates for personal identity, little or no attention is paid to the effect which these parts have on each other. This deficiency is problematic, as the relation between particular ‘person parts’ is crucial to both the identity of those parts, and to the particular way those parts operate.

For example, a particular brain is crucially influenced by the body of which it is part, such that were that body to be different, the brain would be different also. A person with artery blockages, who lives a sedentary and painful life, for example, may develop a different brain structure to, and would have different experiences from a person who is fit and does lots of walking and exercise. While not all differences are as acute as this, they are nevertheless present in some degree or another. This means that the conceptualisation of ‘person parts,' without giving due consideration to the relation between these parts, gives the misleading impression that 'person parts' are independent and self-sufficient. Criterial analysis, however, pays no attention to these anomalies, and considers brains, bodies, and minds as if they were theoretically separable objects, whose role in

personal identity can be determined independently of the role of the other objects to which these parts are connected.

The problem with this approach to personal identity is that it does not accord with the nature of persons as we know them. Persons seem to me to be highly complex entities, who exist in a more-or-less integrated manner, such that it is difficult to conceive how their parts could be theoretically separated in the way suggested by Parfit and similar theorists. I wonder, for example, how could a person's mind be the same, if that person had a different brain? Or how could a person's thoughts be the same, if that person had a different body? Or how could a person's memory be the same, if that person lived in a different place? If we are to really understand personal identity, we must address these questions.

We need to be clear about what persons are, about how their different parts relate to each other, the degree to which these relations are significant, what makes a person change at different times in her life, how the notion of 'change' itself relates to being a person, and so on. As these issues are important, it seems that an appropriate first step to determining the nature of personal identity is to first determine more precisely what persons are. For it is only when we are clear about what being a 'person' means, that we can be clear about what being the *same* person over time means. In the following section, this question is addressed.

1.6 Personhood[19]

Persons appear to be intrinsically complex objects. Pinning down precisely what defines and identifies persons *as* persons is inherently difficult, as the concept 'person' itself is not simple. This is evident from the attempts of many theorists to provide accurate and decisive descriptions of what persons are. This section will look at several accounts of personhood, with a view to locating common

[19]By 'personhood' here I mean no more than the state of being a person. This meaning is intended throughout.

characteristics among these accounts which might help ascertain more precisely what persons are, and what it means to *be* a person.

The concept 'person' proves to have a complex history. The meaning of 'person' has changed as intellectual currents and societies have changed. One thing seems apparent throughout: being a person is being a human being, living in our world, and being recognised by others.[20] The Latin origin of 'person' refers to 'persona' as a 'mask' or covering (Trendelenburg 1910), pp 338-339.[21] Over its history, the meaning of 'person' has expanded to include numerous aspects, including dramatic, legal, social, rational, and moral.

For Locke, the legal and moral aspects of personhood are the most important. The issues of rewards and punishments in this life and the next motivate his concerns. Locke concludes that a person's essential characteristics are intelligence, rationality, and self-consciousness, but does not refer specifically to embodiment.[22] More recently, Peter Strawson claims that being a person implies that one is necessarily an embodied being. He claims that in order to understand ourselves as persons, we must understand others as persons also. For Strawson, this means that we must be able to ascribe psychological predicates as well as bodily predicates to other persons.[23] In order to do this, we must interact with other persons, which means we must be able to distinguish them as being spatio-temporally distinct from ourselves. We do this by means of our separate bodies.

---

[20]While the issue of whether aliens, computers, and animals are persons comes up at times, this issue cannot be taken up here, as it would make the present argument too complex to deal with in the space available. Personal identity is thus dealt with in relation to normal human persons. Once what this involves is established, the question of who or what else might be persons also can then be addressed.

[21]For an account of the very early history of the word 'person,' see (Trendelenburg 1910).

[22]Locke claimed: 'we must consider what *person* stands for:- which I think, is a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places: which it does only by that consciousness, which is inseparable from thinking, and, as it seems to me, essential to it' (Locke 1959), 2.27.11. This claim is taken as foundational to the personal identity debate.

[23]Strawson describes psychological predicates as 'P' predicates, and as attributable only to persons, examples being 'is smiling' or 'is in pain.' He describes physical or material predicates as 'M' predicates, and as applicable to material bodies, whether conscious or not, for example, 'weighs 10 stone' (Strawson 1959), p 104.

So, to be a person is to admit to consciousness, and, at the same time, to admit to embodiment (Strawson 1959), pp 87-116.

Harry Frankfurt conceives personhood to be a special type of agency. He argues that being a person requires the ability to distinguish between first-order desires, second-order desires and second-order volitions. Second-order desires arise from reflection on first-order desires, but are not necessarily intended to motivate action. For example, a physician who, to help his patient (first-order desire), wants to understand what it is like to crave a certain drug (second-order desire), does not want to act on this desire by taking the drug. A second-order volition is a second-order desire which *is* intended to motivate action, such as by wanting to concentrate more on one's work. According to Frankfurt, an entity which can reflect on second-order desires, and distinguish them from second-order volitions, is a person. Entities capable of second-order desires, but not second-order volitions, are *wantons.* Wantons include children, animals and some adults.[24] Frankfurt emphasises that it is the inability for second-order reflection, rather than irrationality which distinguishes wantons from persons. He claims wantons could be completely rational (Frankfurt 1971), pp 9-14.

Daniel Dennett holds that persons have rationality, consciousness, self-consciousness, verbal ability, communicative ability and must be treated as, and be able to reciprocate treatment, as persons. These conditions of personhood are related in various ways. The first three form a set of mutually interdependent conditions. To be rational, one must be conscious, and must be recognised as an object of a 'certain stance' by others. The fourth condition is conditional on the first three, but remains independent of the last two. One might expect that the ability

[24]Although Frankfurt does not specify which adults, I presume he means cases of mental disorder, such as mental illness, dysfunction, retardation, senility, and so on.

for reciprocity requires both language and self-consciousness. But Dennett waves this requirement by considering persons as *intentional systems.*[25]

While the capacity for entities to be intentional systems is a necessary condition of personhood, it is not sufficient. An intentional system does no more than invite simple behaviourist-type responses, such as treating an entity *as if* it had beliefs, desires and so on. This position accommodates the third condition of personhood, but not the fourth. Reciprocity consists of reaction to stimulus, indicating response to that stimulus. While reciprocity, as here understood, seems to presuppose language and self-consciousness (Dennett's fifth and sixth conditions), this is not necessarily so. It is this hiatus in the account which underpins the discord between metaphysical and moral personhood (Dennett 1976), pp 178-181.

The discord is revealed by considering higher order intentional systems, and Rawls' theory of justice.[26] An example of a higher-order intentional system is a second-order intentional system, which has the capacity to reciprocate the behaviourist-type responses of first-order intentional systems. The reciprocity in mind here involves knowing (or at least, predicting) how the other system will respond to a given stimulus. This principle could be iterated to involve higher and higher orders of reciprocation. However, such reciprocation is not restricted to persons. Dennett gives examples, in which other entities, such as animals, demonstrate apparent reciprocity. In one instance, a dog anticipates its master's response by feigning the intention to leave the room, and then retrieving its favourite chair. In another, a bird feigns a broken wing to protect its nest from predators. If these cases are examples of reciprocity, then reciprocity appears not to be conditional on either language or self-consciousness.

---

[25]This concept is brought out in (Dennett 1978), in which Dennett compares three stances which can be taken towards entities, namely the *design* stance, the *physical* stance, and the *intentional* stance. The latter of these assumes rationality, and is thus the best stance to adopt where prediction of behaviour is required. It applies equally to non-human and human entities. See Dennett's article for full details, and for explanations of his reasoning.

[26]As spelt out in full in: (Rawls 1971).

In other words, condition four does not entail conditions five and six. This lack of entailment signals the discord between metaphysical and moral personhood, further elucidated by consideration of Rawls' original position (Dennett 1976), pp 181-185.[27]

The aim of Rawls' position is to formulate principles of justice by initiating responses to given situations. Such responses might be expected to represent a certain moral awareness, but on Dennett's view, just because calculations and deductions take place, does not mean that these responses are not merely part of the 'intentional stance.' Just because reciprocity has occurred, we cannot know for certain that it is accompanied by beliefs constitutive of moral awareness:

> There is no objectively satisfiable sufficient condition for an entity's *really* having beliefs, and as we uncover apparent irrationality under an Intentional interpretation of an entity, our grounds for ascribing any beliefs at all wanes, especially when we have (what we always *can* have in principle) a non-Intentional, mechanistic account of the entity (Dennett 1976), pp 193-194.

In other words, moral awareness is not necessarily coincident with reciprocity. The conditions from which moral awareness could be derived are, in fact, the fifth and sixth. Even then, language and self-consciousness are not of themselves sufficient condition for moral awareness, but are at least necessary conditions of it. Responsibility for an action requires one's awareness of it as one's own action, (derived from self-awareness), which in turn requires one to recognise the action under the relevant description (derived from language). However, while it may appear that this occurs, how can we really know? Although Rawls' original

[27]The original position is proposed by Rawls as the ideal from which to formulate principles of justice. The hypothetical position entails persons evaluating such principles from behind a 'veil of ignorance,' that is, not knowing one's own position or status in society: 'Among the essential features of this situation is that no one know his place in society, his class position or social status, nor does any one know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like' (Rawls 1971), p 12. Rawls' purpose is to locate principles of justice which interpret 'justice as fairness' by building on the social contract theory, as found in the work of Locke, Rousseau and Kant (Rawls 1971), p 11.

position may indicate that moral awareness is present, we can never really be sure if this is the case. The outcome of this view is that not only can moral personhood not be fully secured, but neither, according to Dennett, can metaphysical personhood.

Dennett holds that although we stipulate the conditions of metaphysical personhood, in the end we cannot know whether these conditions are completely satisfied.[28] He claims that personhood is ultimately a normative concept, and that due to the complexities outlined above, is subject to inevitable equivocation and uncertainty (Dennett 1976) , pp 191-194.[29]

Mary Warren stipulates that persons must have consciousness of internal and external objects, reasoning and problem-solving ability, independent self-motivated activity, and self-concepts. Warren's main concern is with the moral and legal status of foetuses, which, although genetically human, have not yet achieved the relevant criteria of personhood. Aliens and robots, on the other hand, could

---

[28]Dennett claims that the concept of a person is 'inescapably normative', and that the 'moral notion of a person and the metaphysical notion of a person are not separate and distinct concepts, but just two different and unstable resting points on the same continuum' (Dennett 1976) , p 193. Michael Goodman agrees with Dennett that metaphysical personhood fails to secure moral personhood, but argues that moral personhood is sufficient, although not necessary, for metaphysical personhood. An example of a metaphysical, but non-moral person is a sociopath, who, while possessing the characteristics of the former, lacks those of the latter. Goodman claims that the crucial condition missing is that of *moral consciousness.* Consciousness that one's actions are either good or bad requires that one is rational, conscious, communicative and so forth, but this moral consciousness is itself additional to those characteristics. See (Goodman 1992).

[29]In later works, Dennett more explicitly denies the existence of consciousness. For example, in (Dennett 1991), Dennett examines in detail analogies between the brain and computers. In taking the view that all phenomena require to be explained from a third-person view-point, he concludes there is no such thing as consciousness. Replying to an 'adversary' who considers how consciousness relates to animals and zombies, Dennett states: 'But why should it matter, you may want to ask, that a creature's desires are thwarted if they aren't conscious desires? I reply: Why would it matter more if they were conscious - especially if consciousness were a property, as some think, that forever eludes investigation?

Why should a "zombie's" crushed hopes matter less than a conscious person's crushed hopes? There is a trick with mirrors that should be exposed and discarded. Consciousness, you say, is what matters, but then you cling to doctrines about consciousness that systematically prevent us from getting any purchase on *why* it matters. Postulating special inner qualities that are not only private and intrinsically valuable, but also unconfirmable and uninvestigable is just obscurantism' (Dennett 1991), p 450; (Searle 1997), p 107. This view has been explored and criticised by many theorists. See, for example, (Searle 1997), in which Searle takes Dennett's view to task, arguing that the experienced qualia of consciousness is undeniable. For a lively exchange between Searle and Dennett on this matter, see especially pp 95-131.

have personhood characteristics and, therefore, could be accepted into the moral community as persons.

Foetuses and defective humans are outside the moral community and therefore lack the corresponding moral rights and responsibilities. This means that the rights of mothers always outweigh the rights of foetuses in cases of potential abortion, regardless of the level of foetal development. Warren argues that because a foetus is always only a potential person, its right to life never outweighs the right to end its life by its mother. This holds equally, regardless of the reason for ending the pregnancy, such as whether the life of the mother is threatened, or whether the mother wishes to take a holiday. Warren holds that although killing a foetus (even a fully-developed one) is not morally wrong, killing a newborn would be. This is because persons other than the parents may want the child, and because people in general prefer to preserve rather than destroy children (Warren 1994), pp 307-311.

Mary Gore Forrester holds that personhood is a moral concept. All persons are entitled to full moral consideration, characterised by the principles of *beneficence* and *fairness.* Whether persons have immortal souls is not an issue for personhood, as persons warrant moral consideration whether they have immortal souls or not.[30] Two types of persons are possible. *Natural* persons are human beings with rationality, consciousness, linguistic competence, self-awareness, the ability to feel pleasure and pain and the capacity for higher-order intentions. *Extended* persons are entities that deserve to be treated with beneficence and fairness, but which lack some of the above characteristics.

[30]This is the view of the present thesis. Whether persons have immortal souls or not, they warrant moral consideration. In addition, as is evident in Chapter 6, whether persons have *immortal souls* is treated, in this thesis, as a different question as to whether they have *empirical selves.*

For Forrester, extended persons include infants, the mentally incompetent, the senile and hypothetical conscious, communicative, rational aliens. Problem cases include foetuses,[31] criminals, and animals. Criminals have violated the principles of justice and fairness, and although they should be appropriately punished, because they are likely to re-enter society, they should not lose personhood. Forrester argues that because animals are sentient creatures, they have interests, and therefore warrant moral consideration, but not necessarily as persons (Forrester 1996), pp 7-13; 78-79; 87-97; 104-113.[32]

1.6.1 Summary of Personhood

The above accounts bring out some of the complexities involved in being a person. For the most part, these accounts focus on mental rather than physical characteristics. Locke points out the importance of consciousness and rationality in being a person. Strawson draws attention to the link between embodiment, consciousness and the ability to relate to other persons. Consciousness and rationality are fundamental to most of the above theorists, although Frankfurt considers the ability for second-order reflection as more primary, and not necessarily tied to rationality. Consciousness itself is problematic, as persons are sometimes unconscious, and are certainly not always self-conscious. Dennett draws attention to the need for reciprocity between persons. His view suggests that unless we accept each other as persons, the whole idea of personhood loses meaning. Personhood is thus a communal affair, and hence requires

---

[31]The question of foetuses is complex, and cannot be given adequate treatment here. Forrester addresses foetuses in *'Should Fetuses be Extended Persons?'* (Forrester 1996), pp 164-187, in which she argues that foetuses are not natural persons, but that a hierarchy of considerations should be brought into play when deciding on their treatment, and under what conditions extended personhood should or should not be granted to them.

[32]Forrester argues that the capacity of animals for reciprocity is limited, and that therefore, to weigh the consideration of animals equally with the consideration of humans would be unfair. Forrester claims that this is not 'specism', but survival. Unless human interests weigh heavier than those of other creatures, the inability to universalise rules would ultimately see humans overrun, and possibly displaced. Additional discussion of animals is presented in *'How Animals Ought to be Treated'* (Forrester 1996), pp 115-128, and *'Animal Rights'* (Forrester 1996), pp 129-136.

Forrester claims that animals should never be treated cruelly, and should be given the degree of consideration commensurate with the resources available in a society. For example, she claims it is never permissible to cause pain to animals in the interest of producing luxury items such as cosmetics, however it may be permissible to eat animals if other food sources are inadequate, and as a result humans would suffer unduly or die out unless animals were eaten.

communicative skills. It is difficult to conceive of a person existing completely by herself, and of never having had interactions with other persons. Interaction between persons seems to be part of what persons are. But, as Dennett also points out, ascertaining whether the characteristics we ascribe to personhood are *really* present may be more difficult that it first seems. For Warren, personhood gives one certain rights over supposed 'non-persons,' such as foetuses. Forrester draws attention to persons unable to engage in reciprocity. She claims that by underpinning personhood with beneficence and fairness, such 'extended' persons could be included in the moral community. These various approaches to personhood demonstrate the difficulty in precisely locating the characteristics of personhood, and, in some instances, in determining whether or not these characteristics are actually present.

The inability to fully pin down personhood characteristics, plus the perceived need by some theorists to make exceptions in problem cases, creates difficulties in specifying criteria for personal identity. It seems that any criterion is open to challenge. If it is too narrow, it excludes the most vulnerable members of our species. If it is too broad, it trivialises personhood. Personhood is thus inevitably a contested concept. There seems to be two important strands of interpretation involved. The first concerns a person's 'natural' endowments, such as embodiment, consciousness, the capacity for rationality and so on. These characteristics appear to be features which most normal humans living in society and interacting with other humans naturally have. The second strand of interpretation concerns the rules of the community in which persons live, as it is these rules that specify which of the 'naturally endowed' characteristics are the most essential. For example, a community that regards rationality as fundamental to personhood might consider that non-rational entities such as foetuses, the mentally handicapped and the senile are non-persons. A community that did not hold rationality to be fundamental to personhood would respond to these non-

rational entities differently, although that community might deny personhood to these entities for different reasons.

For example, foetuses may not be denied personhood due to the absence of rationality, but because they have not reached a certain stage of independence which the community has stipulated. In a community whose resources are extremely depleted, personhood might be denied to the senile in the interests of saving that society from extinction. While rational argument could be produced to justify both these and similar stances, this approach is not without its problems. If the ability to decide who are and who are not persons was totally arbitrary, prejudicial decisions could be made by the strongest members of a community against its weakest members, resulting in the weakest being designated as non-persons.

This discussion demonstrates that specifying criteria for personhood is inherently problematic. Even the providing minimal core criteria is fraught with difficulties. Knowing what to include and what to exclude is no simple task, as whatever is specified is ultimately contestable. Importantly also, features which are specified as essential could be intimately related to other features which are assumed not to be essential. This means that features that are integral to personhood, but which are 'hidden,' are excluded when accounting for personhood, leaving the account deficient and misleading.

As Strawson points out, for example, being conscious (at least in *this* world), requires having a body, as bodies allow us to distinguish ourselves from other persons and to interact with other persons. Also, Frankfurt and Dennett refer to the connection between reasoning abilities and the capacity for action. Just being accepted as a person by other persons requires certain communicative abilities but, as Dennett points out, ascertaining just what these abilities involve is difficult

to discover. What does seem apparent, however, is that many typical personhood characteristics are related to other similar characteristics, and few, if any, appear to be isolated or independent. This is one of the reasons that theorists have difficulty in selecting what seems most essential. When we stipulate 'consciousness,' we may give the impression that consciousness does not carry anything else with it, but, as indicated above, this is not the case. Moreover, even if we do specify certain criteria for consciousness, we cannot assume that we have specified it all. To specify all that is required for an entity to be conscious would require a great deal of knowledge about biology, psychology, evolutionary theory and more. To fully specify what it is to be rational would be equally problematic. Thus, even if some flexibility in criteria is permitted, we can never fully circumscribe what we include. This means that an account of personhood, if it is to be authenticate, should be non-reductive. To fully account for all aspects of personhood would take more than our resources permit. We could never say all that there is to be said.

If personhood is basically irreducible, it follows that personal identity is irreducible also. This means that what maintains personal identity over time cannot be fully specified. We can never know for sure that we have included all that is relevant, and more importantly, whether something crucial, whose influence is present, but also concealed, may have been left out. As the psychological continuity criterion relies on the specification of criteria, it is easy to see why it falls into difficulties. If, for example, we claim that maintaining personal identity over time involves traits, such as rationality or memory retention, we imply many other things as well. When we explicate one characteristic, we inevitably implicate others. We therefore need to rethink the effectiveness of the criterial approach to personal identity. Clearly it tells us some things. But it also misleads us into thinking it can tell us everything that matters. It does not, because it cannot. A different approach to personal identity is, therefore, needed.

1.7 Holistic Persons

Rather than defining persons in terms of discrete characteristics, it is more constructive to consider persons as holistic structures. This way of thinking about persons permits the recognition, not only of the characteristics involved in personhood, but also of the relations between these characteristics. When personhood is 'dissected' into discrete characteristics, the relevance of these relations becomes masked. In most cases, the development and manifestation of particular characteristics is tied to, and interdependent with the development and manifestation of others. To account for one such characteristic is thus to implicate many other characteristics at the same time. Let me outline this new understanding of persons.

Persons are not simple structures, nor are they an assemblage of theoretically separable parts. Persons are complex structures, composed of parts, each of which is indissolubly linked to other parts. Unless the other parts are referred to when making definitions, definition is incomplete. Because the parts are so intimately interrelated, it makes no sense to 'take out' one part, and to assume that the rest are unaffected. The unity of the structure depends on the relation between all of its parts. Indeed, the components are identified and defined only in terms of their participation in the unity. The structure looks 'inward' and 'outward.' It looks 'inward' to its own parts, and to the relations between them. It looks 'outward' as a single structure to its relations with other objects.

We cannot take one part, such as a brain, a limb, a thought, an emotion, or even an environment, away from a person, and assume that neither the removed part, nor the remaining parts are unaltered. Part of what makes that part what it is, is its connection with other parts. Thus, in naming one part of a person, such as a brain, we inevitably name other parts also. It follows that only by considering a person as a whole, can we get some idea of what personal identity involves. If it is

to be authentic, a theory of personal identity needs to recognise this. The remaining chapters take up this challenge, and articulate and defend a theory of personal identity, which conceives of persons as complex and holistic. I will now summarise the present chapter, and then outline the contents of the remaining chapters.

1.8 Complex Persons

The foregoing sections have considered certain metaphysical and ethical implications of the criterial, Reductionist approach of the psychological continuity criterion of personal identity. These implications include the fragmentation over time of personal identity, the blurring of personal identity, and certain ethical ramifications. I claim that these implications are unacceptable, and result from a faulty, atomistic analysis of personal identity. Consideration of the concept 'person' indicates that personal identity is not conducive to reductionist, criterial-type analysis. An alternative, more holistic type of analysis of personal identity is therefore proposed in the following chapters.

Chapter 2 examines some of the thought experiments used by personal identity theorists to argue for the psychological continuity criterion. Thought experiments have been influential throughout the personal identity debate, and have contributed to the atomistic approach taken in contemporary thinking. The scenarios involved consider ways in which the elements that compose personal identity become separated and draw conclusions about which elements retain personal identity. For psychological continuity criterion theorists, these elements are persons' psychological states or minds. By examining thought experiments from several theorists, I argue that such thought experiments are inadequate to comprehensively analyse personal identity. Briefly, they treat the mind-body relation as merely contingent, but produce no evidence for this; their scenarios are often unclear and incomplete; some arguments are unsound, and some

arguments misconstrue the mind's structure. The main purpose of Chapter 2 is to demonstrate the inadequacies of arguments used by many psychological continuity criterion theorists. By drawing attention to these inadequacies, I open the way for further analysis, indicating the direction the analysis should take.

Chapter 3 examines the psychological continuity criterion's internalist view of the mind. This view characterises mental content as internal to the mind. By describing mental content in terms of self-sufficient linear chains, theoretically transferable from one person to another, the psychological continuity criterion fails to acknowledge the mind's external features. No explanation of how mental content is achieved is furnished. As a result, the link between mental content and the events and items in the world from which mental content is formed, is missing from the account of personal identity. This leaves the account inadequate and misleading. In response, I defend an externalist view of the mind. To show how mental content is achieved, I investigate how it is linked to items in the world. My focus is on perception, in particular, on environmental perception and visual perception. Like other biological creatures, humans respond to the environment in relation to their survival needs. Included in this response is the need to secure habitats, and to select those features in the environment which best serve our primary needs. Because response to these features governs and constrains our mental content, acknowledgment of these features is essential to an account of the mind. Visual perception is especially important, as it is a major mode of perceptual response. Consideration of environmental perception in general and visual perception in particular, elucidates the connection between thought and the objects of thought. By recognising the link between perception and the external world, the link between thought and world is acknowledged, and in consequence, the link also between the world and personal identity.

Chapter 4 considers the psychological continuity criterion's characterisation of mental states as atomistic and impersonal. I consider Parfit's view that causal connectedness between mental states could be 'any cause.' While the commitment to Relation R is examined, it is acknowledged that most versions of the psychological continuity criterion characterise mental items atomistically and permit non-normal causal connectedness. To investigate psychological atomism and impersonal description, the psychological continuity criterion's recourse to *quasi-memory* is investigated. 'Impersonal' memory, or 'quasi-memory', is used by the psychological continuity criterion to overcome the criterion's inherent circularity. This strategy results in a defective account of memory, as is shown in an investigation of genuine memory. The investigation brings out the inherently holistic nature of memory, and the role of personal ownership. For memories to have meaning requires that they belong to coherently structured sematic networks, which by definition, are necessarily owned by individual persons with individual histories. The account of memory thus supports the view that an atomistic, impersonal account of memory is inadequate to account for the role of memory in personal identity.

Chapter 5 draws attention to the body's role in personal identity. It considers the psychological continuity criterion's neglect of embodiment when accounting for personal identity. Personal identity over time is taken to be a matter of psychology only. A person's body features only incidentally in the realisation of psychological continuity. Thus, while bodies are needed as a kind of 'vehicle' for psychological continuity, it is irrelevant which persons have which bodies. This denial of bodily relevance takes no account of the body and its functioning, particularly in relation to the brain. The brain and body and parts of the brain and body are treated as if they functioned independently of each other, and as if they played no significant part in mental life. In defending a holistic view of brain and body functioning, I argue that this view is incorrect. The holistic view of the body is supported by

considering the ways in which the body affects mental life, such that were a person to have a different body, that person's mental life and identity would be different also.

Chapter 6 investigates the nature of the self. The psychological continuity holds that no substantial entities underpinning experiences can be found, and therefore, that it is irrational to presume that such entities exist. Persons are held to be no more than bundles of perceptions and it is therefore irrational prefer self-concern over other concern. I argue that this stance on the self is misconceived and results from both an overreaction to Cartesian errors and the failure to recognise that having a self is essential to having experience. While the psychological continuity criterion recognises that experience requires a subject, it does not adequately appreciate what subjectivity involves. I show that subjectivity is more than a peripheral grammatical function and involves a self which is concealed in superficial analysis. I argue that this self is a dynamic unity, which is both necessary to experience and mental life and which is part of what it means to have experience and mental life. I also argue that selves are more than just minds and include bodily capacities for action and the first-person knowledge we have of ourselves and our capacities. In addition, I argue that selves are primarily bodily experienced and only secondarily mentally experienced, and finally, I discuss some aspects of the cognitive self's emergence, and some of the ways in which autobiography and narrative contribute to the development and expansion of the self.

Chapter 7 draws together the main arguments from the previous chapters in support of this thesis' two main claims, namely that:

- Due to its reductionist, criterial approach, the psychological continuity criterion is inadequate to account for personal identity.

- A sound approach to personal identity must respect the complex, dynamic, holistic, non-reductive nature of persons.

This final chapter looks at the implications which the holistic approach to persons and personal identity holds to the ethical problems outlined earlier. It also briefly considers possible future directions that the personal identity debate could take.

Before proceeding further, it should be noted that in arguing for this thesis' claims, great weight is placed on empirical research from a variety of sources. Much of this research is taken from actual persons in actual situations, and is, therefore, more likely to represent personhood and personal identity, than an analysis which relies largely on speculation. Finally, while the claims of this thesis are not intended to furnish a complete account of personal identity, they are intended to indicate some of the important features which such an account should include. The next chapter begins the case for complex persons by investigating thought experiments used by proponents of the psychological continuity criterion to argue for the reductionist conception of personal identity.

Chapter 2 Uncommon Persons

2.1 Introduction

A frequent method of supporting the psychological continuity criterion is the use of thought experiments.[33] These thought experiments concern scenarios, which, although unlikely, appear to be logically possible (Wilkes 1988, p 2. Possible worlds are imagined in which extra-ordinary events occur. Personal identity theorists frequently use thought experiments to inquire into cases where personal identity seems to be problematic. Psychological continuity theorists typically construct scenarios in which mental continuity rather than the bodily continuity is shown to ground personal identity over time. The scenarios concerned usually involve brains or brain parts being transferred into different bodies, or whole persons or parts of persons being duplicated. In these cases, personal identity is presumed to go with the mind, rather than with the body. Persons resulting from these processes have not become persons in the normal way and I thus refer to them as *uncommon* persons.[34] A problem with uncommon persons is that due to their extra-ordinary generation and heritage, we cannot be sure that their personal identity features correlate with the personal identity features of normal persons. Due to this anomaly, I argue in this chapter that the thought experiment strategy which produces uncommon persons is inappropriate to establish definitive conclusions about the personal identity of normal persons.

Philosophy often uses thought experiments to explore obscure or problematic issues. These thought experiments present scenarios in which unlikely events are

---

[33]For a good discussion of thought experiments, see (Wilkes 1988). Some ideas in this chapter are drawn from this work. According to Wilkes, there are different types of thought experiments, those concerning philosophy normally being of scientific, metaphysical, moral, or epistemological interest. Wilkes cites two basic kinds of thought experiments: 1)those which are about scientifically possible situations, such as carrying out a particular manoeuvre to prove a scientific law, as in the case of Galileo contemplating the outcome of objects of different weights falling to the ground. In this case, the procedure was not carried out, but it *could* have been - it was imagined instead; 2) those which amount to playing around with ideas and words in the mind, an example being imagining the grammatical merit of a phrase such as 'colourless green ideas sleep furiously' (Wilkes 1988), pp 2-3.

[34]'Uncommon persons' is my own term, offered here as I believe it suits the situation under discussion.

supposed to take place. Such scenarios present opportunities to speculate outside our normal frameworks. But we need to be careful not to take the 'results' of these experiments too far. We should remember that in most instances, thought experiments describe events which do not usually happen and which may never happen. This should be kept in mind when drawing conclusions from thought experiments, or we could be misled into thinking we have examined something *real* when we have only examined something imaginary. The thought experiment strategy used by psychological continuity criterion theorists is particularly problematic in this regard. It is inclined to draw conclusions on the basis of mere speculation. But, because these thought experiments *are* mere speculation, their conclusions should be taken as indicating where more investigation is required, rather than being taken as having yielded any definitive conclusions in themselves.

Certain typical thought experiments are frequently used by psychological continuity theorists. They usually concern scenarios in which minds are supposedly transferred to different bodies, without personal identity or continuity being compromised. But, as these scenarios are far removed from the norm, we cannot be sure that the events they portray would work out in the same way in real life. In real life, bodily continuity and psychological continuity normally occur together. Between birth and death, persons exist in the world where mind and body operate as a unity. Minds and bodies do not normally become swapped or separated.[35] We cannot be sure, therefore, that minds would not be affected in some significant way if they became 'attached' to different bodies. If the thought experiment strategy of the psychological continuity criterion is to be effective, it must establish that if bodily continuity were to be violated, psychological continuity would remain unaltered. In the relevant thought experiments, however, this issue is not addressed. It is often simply claimed that following the particular disruptive

[35] I am speaking here of normal experiences, excluding situations such as alleged astral travelling, near death experiences, or similar.

processes involved, persons' psychological states would just continue as before, as if this supposed outcome was uncontroversial and uncontestable.

Another problem with thought experiment strategy is that its scenarios must always be incomplete and opaque. This means that details which could be crucial are omitted, leaving important issues obscure and uncertain. In addition, many of the situations portrayed rest on complex and untested assumptions and may not represent genuine possibilities. Key features of the scenarios often deviate widely from the norm and it is often unclear whether these deviations are relevant to the issues under scrutiny. Due to their incompleteness and obscurity, these thought experiments, at least by themselves, are a dubious method of inquiring into the issues which most concern personal identity. In view of these problems, I argue that certain key thought experiments, used by defenders of the psychological continuity criterion, are inadequate to establish definitive conclusions about personal identity. I argue that these thought experiments fail to demonstrate that the mind and body operate discretely. I also argue that they fail to establish that personal identity is grounded in psychological continuity to the exclusion of other forms of continuity. Finally, I argue that they present an inaccurate view of the mind's structure and operation. To argue my case, several key thought experiments are examined. Various responses to these thought experiments are then considered, followed by an assessment of the reasoning used and of the relevance of this reasoning to the personal identity debate.

2.2 Thought Experiments Considered

Locke was one of the earliest theorists to use thought experiments when considering personal identity. A key tenet of Locke's philosophy is that *substance,* both material and immaterial, is unknowable. By contrast, consciousness is knowable, and is the means by which personal identity is preserved. A problem

for Locke is that he considers it possible that consciousness and substance could become separated:

> But yet, to return to the question before us, it must be allowed, that, if the same consciousness (which, as has been shown, is quite a different thing from the same numerical figure or motion in body) can be transferred from one thinking substance to another, it will be possible that two thinking substances may make but one person. For the same consciousness being preserved, whether in the same or different substances, the personal identity is preserved (Locke 1959) , 2.27.13.

But for Locke, where consciousness and soul are conjoined and the body becomes disconnected, it is consciousness that maintains personal identity. To show this, we are asked to imagine that the soul and memories of a prince leave the body of the prince and become manifest in the body of a cobbler. The question of which body then contains the prince permits the more general question about personal identity:

> For should the soul of a prince, carrying with it the consciousness of the prince's past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, every one sees he would be the same *person* with the prince, accountable only for the prince's actions: but who would say it was the same *man*? (Locke 1959), p 457.

According to Locke, we would reject the person with the prince's body as being the prince in favour of the person with the prince's mind. Even though confronted with a different body, with a different history, the mind and memories would have priority as the locus of identity. These claims constitute part of Locke's intended vindication of the psychological continuity criterion.

Inspired by modern neurosurgery, many thought experiments, such as Sydney Shoemaker's, concern surgical procedures in which brain parts are transferred from one person to another. To consider personal identity, Shoemaker asks that we imagine that the brain of one person, Robinson, is removed, and the brain of another person, Brown, is replaced into Robinson's empty skull. If this new 'person' exhibits the characteristics and the supposed memory-knowledge of the

former person Brown, we would be entitled to think that the person Brown now inhabits the body of the former person Robinson, to be now referred to as 'Brownson' (Shoemaker 1984), p 78; (Shoemaker 1984), p 43.

According to this view, Brown has not allegedly become Brownson due to the physical matter of the brain, but rather due to inheriting the required psychological states and to experiencing them in the first-person mode. Shoemaker sees similarities between his brain-transfer experiment and Locke's prince and the cobbler (Shoemaker 1984), p 78. While the physical matter of the brain is relevant as the *vehicle* of transfer, it is only a contingent fact that the relevant psychological states are instantiated in that bit of matter, rather than in any other bit of matter. What is crucial in this case is the inheritance of the right psychological states, as it is these which determine the continuation of identity. It is because it is psychological states that matter, rather than the brain or brain parts, which distinguishes this view as that of the psychological continuity criterion, rather than that of the brain criterion. Recall that the brain criterion focuses on brains and brain parts, rather than on psychological states.

Parfit also uses thought experiments to argue for the psychological continuity criterion. An example of Parfit's thought experiments is one in which a human being is hypothetically teletransported to Mars. Due to imagined advances in technology, it is supposed that it is possible for human persons to become located on Mars without having to travel there in the normal way. At the press of a green button, a brain-scanner is supposedly able to copy and record the complete state of a person's body and brain cells, at the same time at which the body is destroyed. This information is transmitted to Mars, arriving three minutes later. It is then used to create an exact duplicate of the person, using new matter. The person can 'return' to earth in the same way, and in fact, 'travel' back and forth between Mars and earth any number of times. On each occasion, this person

'wakes up' as the self which entered the process three minutes earlier. Parfit imagines that he, himself 'travels' in this way to Mars:

> My replica thinks that he is me, and he seems to remember living my life up to the moment when I pressed the green button. In every other, way, both physically and psychologically, my Replica is just like me. If he returned to Earth, everyone would think that he was me (Parfit 1984), p 200.

Parfit acknowledges that the replica is merely qualitatively similar, rather than numerically identical. But, Parfit argues, according to the 'Widest criterion,' the replica *is* the original person:

> Reconsider the start of my imagined story, where my brain and body are destroyed. The Scanner and the Replicator produce a person who has a new but exactly similar brain and body, and who is psychologically continuous with me as I was when I pressed the green button. The cause of this continuity is, though unusual, reliable. On the Physical Criterion and the Narrow Psychological Criterion, my Replica would *not* be me. On the two Wide Criteria, he *would* be me (Parfit 1984), p 209.

By allowing the causal continuity of mental states to be of any kind, rather than the standard form of memory retention, psychological continuity with the original person is understood to have been maintained. Thus, the person who continues life on Mars is taken to be the same person as the person who formerly lived on earth.

2.3 Thought Experiments Assessed

The above scenarios refer to imaginary cases in which bodily continuity and psychological continuity become disordered or disconnected. In each case, the reader is supposed to conclude that it is psychological continuity, as opposed bodily continuity, which retains personal identity. There are, however, some problems with this approach. For example, the scenarios themselves present their case in such a way that the personal identity concerned is already assumed to be present in the transferred states, even though there are no substantial arguments presented to back this up. We may be inclined to accept this claim at face value,

as we may feel that persons' identities are more likely to be contained in persons' psychologies, rather than in persons' bodies. But in the above scenarios, little exploration is given to what persons' psychologies actually involve. Emphasis is put on the psychology as it was just before the crucial change, but pays little attention to what happens after such change. It is thus assumed that the locus of identity continues as before, and that it is affected neither by the operations, transfers, or changed locations which are supposed to have taken place. Of course, proponents of thought experiments are perfectly entitled to stipulate their own conditions and events, but the problem here is that in doing so, they beg the question at issue. The problem with the above thoughts experiments is that they inquire into personal identity by stipulating the conditions under which personal identity is maintained. This is surely a very odd way to conduct an inquiry.

In taking this approach, the above thought experiments take for granted that psychological continuity, rather than bodily continuity would be favoured as the locus of personal identity. It is not entirely clear, however, just how this conclusion is supposed to be reached, or even could be reached. Would we, for example, necessarily assume that the prince was the person in the cobbler's body just because he acted as if he was the prince or remembered important events in the prince's life? Would this be sufficient evidence that he had the prince's mind? Alternatively, if the prince was an accomplished pianist and the cobbler was not, instead of asking questions about the prince's life, we might ask 'the prince' to prove his identity by playing an item from his repertoire. With his new body, however, the 'prince' might be unable to comply, as he would now lack his former manual dexterity, muscle control, and flexibility. We may then be less inclined to believe that he is really the prince. Personal identity may also difficult to pin down if the supposed prince had the gruff, heavily accented voice of the cobbler, rather than the soft, refined voice formerly associated with the prince. We would be even more confused if the cobbler was a pregnant woman, or if there were large age

differences between the prince and the cobbler. These possible difficulties indicate that personal identity may be more complex than the psychological continuity criterion is prepared to admit.

The other examples could also be problematic if they were expanded a little. We might, for example, question the identity of Brownson if, in addition to mental characteristics, we also consider physical ones. We may not necessarily agree that Brownson was the former Brown if Brown was ethnically different to Robinson - say Brown was Chinese, short and small-framed, and Robinson was Jamaican, tall and muscular. We might then be inclined to say that neither psychology alone, nor brain identity are sufficient to capture personal identity. While this is an extreme example, it draws attention to the fact that identity-determining features are not captured simply by a single brain or a single set of psychological states, and are not, therefore, as clear-cut as the psychological continuity criterion makes out. Parfit's case could also be confusing. For example, what if the duplicating machine ran amuck, and produced hundreds of Parfits, and scattered them all over the universe, would they all be Parfit, or only one of them? If none of them were Parfit, but then all but one of them died the following week, would the one which remained then be Parfit?

These kinds of problems and questions suggest that thought experiments provide a fairly superficial analysis of personal identity. Those above fail to grapple with many pertinent issues which could be involved. The work of a number of other theorists also supports the idea that thought experiments of this sort are indeed highly problematic. For example, Bernard Williams mounts an objection to Locke's case by presenting a similar thought-experiment.[36] He asks us to imagine that Charles wakes up one day with memories which appear to be those of Guy

[36]For this objection, see (Williams 1973), pp 1-18. The fact that certain thought experiments throw doubt on other thought experiments does not undermine the claim that thought experiment strategy is inadequate on its own to account for personal identity.

Fawkes. Charles remembers having done certain things which earlier he could not remember doing, and cannot remember doing certain other things which earlier he could remember doing. Even though these new memories appear to be first-person memories, the strangeness of the situation might prompt a listener to Charles' tale to try to verify that it was Charles who performed the deeds in question. For Charles' testimony to be valid, the listener might seek a witness to verify that Charles was bodily present at the events concerned. Williams argues that if there was no such witness available, the situation could be addressed from a different angle. Rather than first individuating Charles as an agent, and then ascribing a specific action to him, a particular action could first be individuated, and then uniquely ascribed to an agent, for example, 'the person who murdered the Duchess, whoever it was.' Under this latter approach, should Charles' actions prove to be those generally understood to have been undertaken by Guy Fawkes, we may be inclined to believe that somehow Charles has become Guy Fawkes (Williams 1973), pp 4-8.

Williams claims that even if we could overcome obvious objections to this account, such as our incredulity at the idea of re-incarnation, or the apparent difference of personal and bodily characteristics between Charles and Guy Fawkes, there is one particular reason why this account should not be accepted. What if Charles' brother Robert also makes the same claim, that is, he also claims to have witnessed and carried out the relevant actions? In this case, they might both be, not only Guy Fawkes, but also each other. Williams claims this outcome is 'absurd.' Yet, if Charles and Robert were equally good candidates for the identity in question, there would be no principle to determine *which* of them was Guy Fawkes. Williams claims that the analysis undertaken here fails, due the failure to include the body:

> We are trying to prise apart 'bodily' and 'mental' criteria; but we find that the normal operation of one 'mental' criterion involves the 'bodily' one (Williams 1973), p 5.

By 'separating' the mind from the body, we have ignored the point that part of what it means to remember oneself performing a particular action, is to have the body which performed that action. This point creates serious problems for Locke's approach and for those derived from it. Unless we recognise that particular bodies ground particular sets of psychological states, there seems nothing to prevent the indefinite proliferation of single set of psychological states. The relation between the body and personal identity needs to be taken more seriously.[37]

Person reduplication is not only problematic for psychological states, but also for the brain and the brain parts in which they are instantiated. An early version of this problem is raised by David Wiggins.[38] Based on Shoemaker's account of Brown and Robinson, Wiggins considers the possibility of duplicate persons.[39] Wiggins extends the scenario to Brown's brain being split into two equal halves prior to the transplant. Each brain-half is put into a different body, resulting in two persons with the former Brown's memories. Based on the psychological continuity criterion, both new persons have equal claims to now being Brown. This outcome entails the unacceptable consequence that they will each initially appear to be the same person as each other, but as they each take up threads of different lives, they will chart separate histories, and then 'become' two different persons:

> if we say each is the same person as Brown, we shall have to say Brown 1 is the same person as Brown 2. That is an inescapable part of what was meant by saying that each was the same person as Brown. But Brown 1 will have all sorts of experiences which Brown 2 will not. They will be in different places and have separate experience from now on. And they will communicate *interpersonally* (Wiggins 1967), p 53.

---

[37]Williams' solution was later criticised, on the grounds that bodies also can be duplicated. This has, of course, meant that further thinking is required to overcome the problem of person reduplication (fission), but Williams' work has nevertheless opened up the issue to discussion.

[38] Another seminal article in this debate. See (Wiggins 1967), pp 50-58.

[39] Wiggins takes his discussion of this example from Shoemaker's presentation of the case in (Shoemaker 1963), pp 23-24, see (Wiggins 1967), p 50 & p 78. I use the example in Shoemaker's abbreviated version of the same scenario in (Shoemaker 1984) , in which Shoemaker (on p 40) refers to Wiggins' variation.

This outcome is rife with paradox and confusion. Is Brown one person in two bodies, two different persons with two different lives, or does he cease to exist altogether? The consequences of brain-body discontinuity overturns our common-sense notions of personal identity. These notions are further disturbed when additional brain dissections and transfers are considered, as there then seems to be several persons where previously there was only one.

Parfit's case of imaginary teletransportation is also subject to problematic interpretation. This scenario is supposed to justify the 'Widest' criterion - psychological continuity with any cause. Based on this criterion, causal connections of any kind could exist between the psychological states of an individual at one time and the identically similar states of another individual at a later time. This argument requires that the relation that counts is the survival of those states, rather than the survival of their owner. Two objections to this argument are considered. Both objections attack Parfit's causal requirements. Andrew Brennan claims that Parfit's causal requirements imply that the survival of psychological states is equally as indeterminate as is personal identity (Brennan 1987). Robert Elliot claims that causal requirements do not entail the survival relation, are virtually meaningless, and could therefore be dispensed with (Elliot 1991).

Parfit's claim is that the 'Widest' criterion shows that what matters is the survival of particular mental states, regardless of their particular physical embodiment or of the cause of their survival. Brennan's claim is that the 'Widest' criterion shows that both survival and personal identity are equally indeterminate (Brennan 1987), pp 225-230. The problem seems to be in the apparent relation between survival and personal identity. Two aspects of Parfit's theory are at issue: the need for a causal connection (of whatever kind) for the survival of mental states, and the irrelevancy of whether the person who causes the states survives or not. Brennan gives the

example of the mental states of one individual 'surviving' in a different individual. If I strongly influence my friend, my mental states (as they were at the time of my influence) may survive in my friend, whereas, due to my changing ideas, such states may not survive in me. There is thus less difference between my past mental states and the present mental states of my friend, than there is between my past mental states and the present states mental of myself. In this case, I have 'caused' the survival of certain of my past mental states in my friend, but this fact is of no interest to me. I am more concerned with the survival of those mental states that concern me now. Thus, whether mental states are causally connected is not the issue which determines their level of importance to the individuals who have them. Yet, on Parfit's account, causal connectedness is *all* that is important. Causal connections are more important that persons.

According to the 'Widest' criterion, a relation exists between the survival of mental states and the individual who causes them. Yet, on this criterion, the survival of particular individual persons is unimportant. This implies that their survival is in some way distinct from, or independent of the mental states they cause. But if their survival is independent of those states, they cannot be the cause of them. Alternatively, if particular persons *are* relevant to the cause of mental states and their survival is unimportant, then those states are similarly unimportant. Put another way, if having a cause matters, then so must the person who causes - that person *is* the cause. If the person had not survived, the states in question would not have been caused at all. But if the identity of the person is merely trivial, then so also must be the causal role attributable to that person:

> If causal differences of the sort mentioned are trivial, then either personal survival is independent of causal role in the way just suggested, or, if it does depend on causal role, it depends on merely trivial circumstances and is thus - given Parfit's view - no better off than personal identity (Brennan 1987), p 226.

Brennan claims, however, that if the causal role *were* abandoned, there would be an unacceptable proliferation of mental states. Without causal connections to anchor them to a legitimate origin, any selection of mental states deemed to be sufficiently like the originals could count as cases of survival. And without the refinements attributable to particular individuals, there would be 'too many cases to count as cases of survival.' Brennan concludes that unless the role of persons as causes is clarified, arguments about survival are subject to 'crippling ambiguity' (Brennan 1987), pp 226-230.

Elliot argues that Parfit's notion of cause lacks the elements required for the survival of mental states. He notes that Parfit's account of psychological continuity requires a causal connectedness condition if the psychological states concerned are to be successive states within the same chain. But he also argues that Parfit's causal continuity requirements (CCR) are ineffectual and do not contribute anything worthwhile to Parfit's theory. If CCR was dropped from Parfit's theory, however, the theory would not work, as there would be no stipulated requirement for causal connectedness among psychological states and thus no explanation of how psychological continuity was maintained. On the other hand, even if CCR was retained in Parfit's theory, if it is ineffectual, its retention would be virtually meaningless:

> My concern is rather to show that psychological continuity theories which include CCR are unstable; either CCR must be dropped or the psychological continuity approach must be abandoned (Elliot 1991), p 58.

Elliot uses two sets of thought experiments to argue his case. The first in each set are cases where the presence of CCR seemingly warrants the conclusion that identity is retained. The second in each set are cases where the absence of CCR seemingly implies that identity is not retained. Elliott argues that the difference in the cases is not sufficient to warrant different conclusions in each set of cases. He

argues that if identity is maintained in the first examples, it should also be retained in the second.

In case *Aa,* a super-being creates Y following the death of X. The super-being creates Y to be psychologically similar to X because it wants X to live on in Y. Both X and Y know this, and expect it to happen. No bodily or brain continuity is involved, but there is a causal connection in virtue of the super-being's specific intentions and actions. Based on CCR, we should accept that X lives on as Y. Case *Ab* is similar to *Aa,* in that Y appears following the death of X, complete with similar psychological states to X, and these events were expected by them both. However, there is no causal connection between the death of X and the appearance of Y. The fact that no such connection exists does not affect the fact that the psychological states in Y happen to be similar to those of X. Elliot claims that if we accept that X survives as Y in case *Aa,* we should also accept it in case *Ab.* The reason is that although a causal relation is present is *Aa,* that causal relation is not one of survival. It is no more the case that X caused the states of Y in *Aa* than it is in *Ab.* Thus, if we are prepared to accept X as being Y in *Aa,* we have no legitimate grounds for denying it in *Ab.*

The second set of thought experiments is similar to Parfit's teletransportation case. In *Ba,* X becomes located at a distant place as Y, by having her body biochemically recorded, reconstituted as Y, and then destroyed. CCR is met due to the causal connection between the various stages of the process. In this case, we would, in accordance with CCR, accept that Y was the former X. Case *Bb* is similar, except that the blueprint containing X's records is lost. During the malfunction, a person Y appears. She has been constituted from a stockpile of elements, but is coincidently just like X. Further, Y believes that she is X. Things for both X and Y seem to have occurred just as they had expected. Thus, even if Y discovers the malfunction, she still believes she is X. But although X mirrors Y,

no causal connection between them exists. Elliot claims that in spite of this, if we accept that Y is X in *Ba,* we should also accept this in *Bb.* This is because, although CCR was present in *Ba* and not in *Bb,* the causal connection was not one in which the states of the original person actually survived. In essence, the person Y in *Ba* is no different to the person Y in *Bb.* This means that if the psychological continuity of *Ba* is sufficient to satisfy CCR, then so also is the psychological similarity of *Bb*. The conclusions drawn from both sets of thought experiments is that Parfit's version of causal continuity does not capture the elements involved in the survival of mental states. Consequently, psychological continuity is rendered virtually meaningless, as there is no discernible difference between instances when it pertains, and instances when it does not.

2.3.1 Summary of Assessment

In reconsidering the above thought experiments, a number of deficiencies have been brought out. It appears that determining personal identity on the basis of psychology alone is not clear-cut and that other features, such as physical or bodily characteristics could also be involved. In Locke's example, it was claimed that in virtue of having the prince's memories, we would take the former cobbler to be the prince. Similarly, for Shoemaker, due to the brain transfer, we would assume that Brownson was the former Brown. Finally, the precisely controlled copying and duplication process encouraged us to take Parfit's Martian duplicate as being Parfit. But it was pointed out that if we took into additional factors, such as appearance, physical characteristics, or possible technical malfunctions, these supposed conclusions could be undermined.

Williams and Wiggins show how some thought experiments can be altered to produce peculiar results. By extending the cases of inherited memories, or by increasing the amount of dissection carried out on a single brain, multiple persons could supposedly be produced from a single person, raising the bizarre spectre of possible uncontrolled person proliferation. Brennan and Elliot reveal flaws in the

causal continuity requirements of the 'Widest' criterion. They show that the refusal to stipulate conditions which distinguish genuine causes from merely superficial ones means that 'causal continuity with any cause' ultimately amounts to no cause at all. Thus, although certain conclusions about personal identity are supposed to flow from thought experiments, it is apparent that these conclusions are in no way definitive, as they are extremely speculative and are always subject to modifications which could produce different or even opposing results. We can never be sure which thought experiments to take notice of, or indeed, whether to take notice of any of them at all.

2.4 Locating the Difficulties

Although thought experiments have long been used as a tool of philosophical inquiry, we need to be aware of their limitations. They are, after all, merely a tool, to be used in conjunction with other available tools. If we take them too much at face value, we are in danger of accepting the purely fantastical as being more important than the actual. This appears to have happened in the case of the those considered above. They treat the issues involved in personal identity in a very simplistic way, such that a particular answer as to 'wherein lies the identity of the original person?' is supposed to be obvious and uncontroversial. We are wooed by the simple elements of the stories to presume all too easily what the answer might be. To accurately assess the value of these stories, we need to consider them in terms of the actual world, rather than just in terms of imaginary worlds.

Wilkes notes that many thought experiments lack specific background conditions. These would be essential in scientific experiments, as they affect the legitimacy of the experiments' results. The thought experiments under discussion are not commensurate with scientific experiments, so they cannot carry an equivalent weight. Wilkes also points out an important difference between the use of thought experiments in philosophy and in the use of fantasy in literature. In the case of

fantasy, an environment is supplied in which fantastical events can occur, permitting our suspension of belief. The world of Carroll's Alice is one in which it is legitimate to abrogate the laws of nature; we know her world is not intended to be one commensurate with our own:

> A world in which one can walk through mirrors is, as explicitly indicated, a world of a dream; in such a world mushrooms can make one grow or shrink, a shop can turn into a boat, Queens can believe six impossible things before breakfast. For such fantasy, we have another world sketched for us, against the background of which the events are intelligible (Wilkes 1988), p 10.

In the thought experiments under discussion, however, such background conditions are not sketched. It is not clear whether we are operating under our own laws of nature, completely different laws, or even laws contrary to our own. If the world of the foregoing scenarios was one in which persons did inherit the mental states of others, or in which it was possible to transfer brains, or in which persons could be reduplicated, would there be other changes entailed by these occurrences which would affect our judgments about personal identity? The situations presented may not only be scientifically or metaphysically dubious, they may also be logically inconsistent. As Wilkes points out, a world in which gold has a different atomic number, or where water is no longer H_2O, is arguably as impossible as one in which a fish could be a whale (Wilkes 1988), p 18. In other words, if we reinterpret one concept, we need to recognise the impact this could have on the concepts to which it is related. If in our world the body and brain are required for our mental states, we need to consider whether it is intelligible to assume we can radically change the body and brain, without our mental states also being radically affected.

These points are relevant if we are to draw strong conclusions based on theoretical brain transfers or person reduplication. It matters whether or not the difference between these imagined cases and the cases of real life are significant

differences. For example, even if we lack the skill to successfully separate brains from bodies (at present anyway), does the theoretical separation make any sense? Brains are not inherently discrete objects, but are part of the nervous system, a complex network which exists throughout the whole body. Would the laws of nature be different in a world where transferring such a large and fragile system was possible? Similarly, even on a purely materialist interpretation, does the idea of person replication make sense? Persons are not static objects, but from the level of cellular activity to the life-preserving functions of the body's major systems, are subject to constant dynamic change. At the sub-atomic level, not all the features of particles can be determinately specified. There are genuine doubts that 'copying' a person makes any sense. As mentioned earlier, we need to remember that on the one hand, while thought experiments are perfectly entitled to stipulate what conditions and events they like, on the other, we need to be aware whether in doing so they simply beg the question at issue.

These considerations raise two crucial issues. First, do the thought experiments in question address the key issues involved in personal identity? Second, do they yield the conclusions claimed? In other words, are the issues raised in these scenarios the right ones, and do the conclusions claimed follow from the premises? Let us grant for the moment that the various forms of mental content inheritance, memory transplant, and reduplication are possible. In what way do these scenarios actually show that personal identity is a matter of psychological continuity, to the exclusion of other forms of continuity? Consider the following scenario:

The earth has been devastated by nuclear winter, resulting from a massive nuclear war. As a result, individuals have a much shorter life-span, only twenty years as adults. Technology is such that old bodies can be recycled and re-constituted into new bodies. Individuals expecting to die can thus order a new

body just like their own ahead of time. When death is impending, brain states can be copied, such that the resulting individual continues life much the same as before. In order to keep life as normal as possible, individuals voluntarily agree to undergo no more than three such transplants in their "life-time," so that "death" finally occurs at about age eighty years.[40]

According to the psychological continuity criterion, because they have a continuing psychology with the right kind of cause, the persons above remain the same persons over time in virtue of the continuation of their psychological states, rather than because of the continuation of their bodies. The elements in the story are sufficiently similar to those of the earlier stories to recognise the points at issue. Let us now pursue the two questions outlined above.

First, have the issues involved in personal identity been addressed? We could answer in the affirmative, on the basis that we have referred to the continuation of individuals' psychology and individuals' bodies. Old bodies die, while psychology lives on in new bodies. Based on this, personal identity is claimed to be preserved because psychology is preserved. But this answer does not show that the full range of issues involved in personal identity has been addressed. It just shows that we have raised the issues we intend to address. That is, we have explored personal identity by addressing psychological continuity and bodily continuity. We have begged the question whether the issues involved in personal identity are confined to psychological continuity and bodily continuity. In other words, for the above thought experiment to work, we must already have decided what is involved in the question of personal identity. The thought experiment itself did not reveal anything we did not already 'know.' It did not explore other issues which might be involved, such as the environment, relations, whether a self is involved, and so on.

---

[40]This thought experiment is my own, although I may have read one like it somewhere, but cannot be sure.

The above approach to personal identity is the same as that taken by the previous thought experiments. In those discussed, we are given a set of factors pertaining to personal identity, and asked to choose amongst them, for example the prince's memories, the prince's body, the cobbler's body, Brown's memories, Robinson's body, Parfit's duplicate, and so on. Because we are presented with a set of factors, amongst which a choice of alternatives is possible, it is tacitly assumed that the total set of factors has been presented. But it is not at all clear that this is the case. If factors other than those presented were involved in personal identity, we would not learn this from the thought experiments themselves. This means that by confining our analysis to what *is* involved in the thought experiments, we are giving tacit approval to the range of possibilities that they provide. We are thus lured into thinking we have given personal identity a comprehensive analysis, when in fact we have not.

Second, we need to consider whether the above thought experiments yield the conclusions claimed. In the example just given, we are to assume that individuals remain the same persons over time in virtue of retaining appropriate psychological continuity, in spite of periodical body replacement. But in what way does the scenario show that personal identity is a matter of psychological continuity rather than bodily continuity, or indeed, any other form of continuity? How has the story proved that it is the fact of psychological continuity alone which preserves identity, rather than say, social custom, community acceptance, stability of relationships, the existence of a self and so forth. There is no specific argument to show that it is psychological continuity over and above these other things which encapsulates personal identity. To accept the case, I must already be committed to that view. In other words, unless I *already* believe that personal identity is a matter of psychological continuity, I will not be convinced by a story which merely reiterates that view, without arguing for it. To conclude from these thought experiments that persons retain their identity in virtue of psychological continuity, in spite of various

bodily changes, whether due to swapping of minds and bodies, brains and bodies, or complete reduplication of bodies, and so on, one must *already* be sympathetic to the primacy of psychological continuity. The thought experiments themselves do not argue for this, they merely demonstrate 'cases' of it. This point could be put in the following way:

Personal identity is a matter of psychological continuity. We can see this because in cases where bodily continuity and psychological continuity become disconnected, personal identity is maintained in virtue of psychological continuity. That is, we can see that personal identity is maintained because we can see that psychological continuity is maintained. Therefore, personal identity is a matter of psychological continuity.

This argument does not justify the primacy of psychological continuity. It demonstrates neither that psychological states operate independently of physical states, nor that personal identity does not involve issues other than mental continuity. If other issues were involved, they would not be revealed in these types of arguments.

A final problem is a serious flaw in the general strategy of the above thought experiments. The flaw concerns the type of analysis which is given to mental states. In the various scenarios, mental states are presented as *insular,* as self-sufficient and self-contained. That is, they are considered in terms of their intrinsic content. This view, sometimes known as *internalism,* stipulates that what makes a mental state the particular state it is, is describable in terms of the mental state itself, as opposed to something external to the state, such as some item in the person's environment. For example, in the case of Locke, what makes the former cobbler the prince is that he has the former prince's mind contents. It is not altogether clear just what this means. Perhaps it involves having mental pictures of the palace where the prince lives, or of his courtiers, jewels, and fine clothes. If

the cobbler lives in a crude village setting, it is not clear how he would recognise these mental pictures. Similar questions could be asked of any thought experiments which discuss mind contents as if they are items contained wholly inside a mind or a brain. Such internalism is controversial and highly questionable, and will thus be addressed in more detail in the next chapter.

2.5 Summary

This chapter has evaluated the efficacy of key thought experiments in the personal identity debate. Investigation of specific examples has revealed several inadequacies. First, in discussing the various forms of mind-body separation, mind and body are presumed to operate discretely, without adequate grounds being produced that this is the case. When describing various scenarios, in which minds and bodies, or brains and bodies become separated, the relevance of connections between these features is not examined. Whether embodiment affects mind content is not even questioned. Second, the thought experiments are sketchy and incomplete. Without the specification of background conditions, it is unclear whether the omitted details are relevant to the case being argued. While the laws of nature under which the scenarios are sketched are not specified, the scenarios themselves do not appear to cohere with the laws with which we are familiar. Third, the lack of specified conditions and the incomplete arguments render claims inconclusive and open to counter-examples and counter-arguments. Through lack of detail, potential logical inconsistencies and impractical or contradictory states of affairs are not revealed. Fourth, the thought experiments do not offer cogent arguments in support of the psychological continuity criterion, but operate as if the case for it had already been proven. Finally, the thought experiments adopt an internalist view of mental states, such that no explanation of the meaning of those states is furnished, other than that of internal reference to the state itself. It will be argued that this conception of mental states is ultimately unsound and cannot be sustained.

In summary, these key thought experiments fail to vindicate the psychological continuity criterion, and ultimately leave the issue of personal identity unresolved. What the thought experiments do reveal, however, is that the analysis of personal identity requires a broader investigation than pure speculation alone can provide. Some areas in particular need further inquiry. these include at least, the structure of mental content, the mind-body relation, and also a more developed and realistic inquiry into the nature of the self. As a first step in this inquiry, the psychological continuity criterion's internalist conception of the mind is taken up in the next chapter.

Chapter 3 Externalised Minds

3.1 Introduction

The main tenet of the psychological continuity criterion is that personal identity is dependent on a person's psychological states or mind. Diachronic identity is taken to be maintained due to the continuance of mental content with the right kind of cause, which for Parfit, could be any cause. The significance of mental content lies in its causal connectedness, rather than in the content itself. As long as causal connectedness is maintained in some form or another, personal identity is also maintained. In cases where brains and minds are supposedly transferred to different bodies, the identity of the resulting person can be determined by checking to see if causal connectedness is present between the earlier person and the later person. Checking whether Charles is Guy Fawkes, Brown is Brownson, or the Martian duplicate is Parfit, is just a matter of ascertaining whether causal connectedness is present between the minds of Charles and Guy Fawkes, the minds of Brown and Brownson, and the minds of Parfit and his duplicate. The actual contents of these “persons' “ minds is of less significance than the causal connectedness - of any kind - which is present between the items which make up those contents. The problem here is that in permitting causal connectedness of any kind, this view takes no account of the factors external to the mind from which those contents first arise. This means that there is no correlation between what goes on inside the mind, and what goes on outside the mind. In this chapter, I argue that this internalist approach to the mind and its contents is mistaken, as it represents neither how minds function, nor how they relate to personal identity.

The internalist conception of mind holds that minds are autonomous and self-contained.[41] Mental content is not held to depend in any essential way on anything existing external to the mind. Minds are entirely internally constituted, and identifiable solely in terms of their intrinsic content. This means that the actual world in which thinking subjects are situated is not taken into account when giving an account of the mind. The opposing view is that minds are interdependent with the external world and that reference to the external world is required to identify and individuate mental contents.[42] The distinction between these two views is broadly referred to as that between *internalism* and *externalism*.

The internalist conception of the mind is necessary to the viability of the psychological continuity criterion. For, if external factors were acknowledged to be significant in determining mental content, then the psychological continuity criterion would not adequately capture what is essential in personal identity. More strongly, if external factors were shown to be *essentially* involved in the formation and maintenance of mental content, the psychological continuity criterion would be seriously undermined. The internalist view of the mind, however, is problematic. Investigation of mental content shows that what happens 'outside' the mind is crucially involved in what happens 'inside' the mind. In what follows, I will argue that the internalist view of the mind is inadequate to address to the actual causes of mental content and thus cannot adequately account for psychological continuity, or for personal identity. Although causal connectedness is stipulated by the psychological continuity criterion, the causal continuity requirements effectively preclude the actual causes of mental content from being recognised. Only a view which acknowledges the mind's external factors can recognise these causes. Therefore, I also argue that only an externalist view of

[41]Descartes, Leibniz, Berkeley and Hume are examples of thinkers who hold an internalist view of the mind and its contents. See, for example: (Descartes 1968); (Leibniz 1973); (Berkeley 1965); and (Hume 1888).

[42]Later Wittgenstein, Putnam, and Davidson are examples of thinkers who hold an externalist view of the mind. See, for example: (Wittgenstein 1958); (Putnam 1981); and (Davidson 1995).

the mind can adequately account for psychological continuity or for personal identity. These claims will be defended, first, by explicating the relevance of the internalist-externalist distinction to the psychological continuity criterion, and second, by showing how empirical research supports the externalist case. This research focuses on perception, in relation to both perception of the environment in general, and visual perception in particular.

3.2 The Problem with Internalism

The internalist view of the mind focuses on factors 'inside' the mind, to the exclusion of factors 'outside' the mind. When accounting for mental content, the internalist view does not make explicit reference to the objects[43] which first caused that content. When mental content is considered more closely, however, it becomes apparent that content *inside* the mind arises, first and foremost, on the basis of objects *outside* the mind. Investigation shows that mental content is dependent on external objects existing. The types of minds we have, as well as the contents of our minds, is related to objects which exist outside the mind. This claim does not just refer to the *types* of minds and mental content we have, but to the fact that we have minds *at all.* If external objects are crucial to minds, a comprehensive account of the mind must include reference to these objects and to our mode of involvement with them. The psychological continuity criterion, however, is remiss in this regard. When giving an account of the mind, the psychological continuity criterion does not make explicit reference to the objects which furnish the material for mental content. This is because 'any cause' of mental content is allowable, meaning the actual cause is often overlooked.

The deficiency in the psychological continuity's approach to the mind can be understood by considering more closely the factors on the basis of which mental

[43]'Object' in this chapter, signifies event, fact, object, or place, in fact anything that humans perceive, engage with or think about.

content emerges. Mental content emerges from our interactions with objects in the world. Without these objects and interactions, mental content would not be possible. This means that when accounting for the mind, these fundamental causes of mental content should be specified. If they are not specified, the existence of objects fails to be acknowledged as a necessary condition of thoughts about objects. This implies that thoughts about objects are prior to the objects themselves, and that the very intelligibility of thought does not depend on objects actually existing.

When considered closely, the internalist view of the mind is paradoxical. If the objects which cause thought are left out of the account of thought, a theoretical division is thereby established between the thoughts of objects, and the objects of thought. More specifically, if the psychological continuity criterion does not acknowledge the *existence* of the objects as a precondition of *thoughts* about objects, there is an explanatory gap in its account of mental states, as it has not provided an explanation for the *content* of thought.[44] But if the psychological continuity criterion does provide this explanation, its ability to comprehensively characterise mental states independently of non-mental factors is undermined.

The weakness of internalism can be seen by considering different kinds of thoughts. These thoughts could be about general items, particular items, and *this* particular item here and now, such as the pen in my hand. First, I can think about cats in general, without having to think about one particular cat. But unless I know about cats, or have engaged with cats, I could not have any thoughts about cats. Indeed, unless there were cats in existence, *no-one* could have thoughts about

[44]It is true that fictions and fantasies are possible about any object. But they could not get off the ground unless there was some connection, either direct or indirect, with something which already exists. This same point applies when mistakes in thinking are made. Even in cases of misremembering or misidentifying objects of thought, connection to something which exists is essential. Even in mistakes, the *content* of thought could not arise totally ex nihilo. If no objects existed at all, we would have no starting point for our thoughts.

cats.[45] Second, I can think about a particular item which I have never seen, or engaged with directly, such as Mount Everest.[46] I can do this, although I have never visited Mount Everest. Even had I never climbed a mountain or a steep hill, I can still have thoughts about Mount Everest. I can do so because I have learned about Mount Everest from people who *have* visited it, or who have had either direct or indirect engagement with it. Persons who have engaged with Mount Everest could not do so unless it existed. Genuine thoughts about Mount Everest are thus contingent on Mount Everest actually existing. Finally, I can think about *this* particular pen here, now, even when I am not looking at it, or using it. But I can only do so because I *have* looked at it, and *have* used it. My engagement with the pen, and my knowledge of the pen is first-person knowledge, and is contingent on the pen being present to me as a physical object. As a physical object, it has interacted with *my* limbs, and *my* sense organs, such that my thoughts about it are directly related to this interaction. In all the above cases, the thoughts about objects are contingent on, and arise out of, the objects actually existing.

The above examples indicate that there is a necessary connection between the objects of thought and the world of objects, or put another way, between what we can think about and what exists. The crucial point is that for thought to occur *at all* requires that objects exist. Our whole understanding of 'what a thought is' is tied to the notion of objects existing in a world. If no objects existed, no thought would be possible. You only have to imagine yourself living in a world (a world? what could this mean here?) with no other objects of any kind beside yourself. What

---

[45]If cats did not exist, the thoughts we have about cats would not be about the cats that *do* exist - they would be about imaginary creatures. While it is possible to have thoughts about objects thought not to exist, such as unicorns and mermaids, this does not undermine the point being made. Imaginary objects can only be thought about due to their connection with objects which *do* exist. This point is essentially the same as that made above. If we *never* interacted with objects, we could not have thoughts about real objects *or* imaginary objects.

[46]Again, if Mount Everest did not exist, my thoughts would not be about Mount Everest, they would be about some other imaginary mountain. You may think: 'Well, how do you know it really exists, couldn't all the stories you read be mistaken, or a trick'. Of course, this is possible in some instances, but it could not be possible about *all objects.* For such fictions to be possible, *some objects* must exist in the first place.

could you think about? Indeed, how could you even know what you were? In such a case, even *a priori* thoughts would not be possible. *A priori* thoughts are only possible for beings who have minds. And beings who have minds are beings who exist in, and interact with, a world of objects. If the very notion of thought itself is contingent on the world existing, a comprehensive and accurate account of thought should include reference to that world. The psychological continuity criterion fails in this regard. When accounting for mental states, it does not require that the objects from which the thoughts, memories, and other mental items first arise, are explicitly specified in such an account.

In holding this position, the psychological continuity criterion takes the connection between thought and the objects of thought for granted. It assumes that because thought and the objects of thought can be disconnected in the imagination, the existence of objects is not required as a precondition of thought. This is misleading. Although the psychological continuity criterion does not directly dispute the connection between thought and objects, it does not emphasise this connection. As a result, it draws inaccurate conclusions about mental states and consequently about personal identity. More specifically, because the psychological continuity criterion fails to recognise the necessary connection between thought and the objects of thought, it is able to execute the theoretical separation between thought and world, and consequently, to categorise psychological states independently from the items which initially caused those states. As a result, the account of personal identity, produced by the psychological continuity criterion, fails to include many significant items, such as bodies, places, relationships, activities, events, and in fact, many of the items which characterise and personalise individuals' lives. In failing to include these items, the account of personal identity is incomplete.

3.3 The Case for Externalism

In contrast to internalism, externalism is the general view that mental content is constituted, at least in part, by factors external to the mind (Honderich Ted 1995), p 265. This view recognises the interdependence between the contents of minds and the world of objects. Work by Donald Davidson and Jeff Malpas helps to elucidate this view.[47] Davidson's work, in particular, is especially important, if only because he is one of the most influential figures in the formulation of a version of the externalist position.

Davidson recognises the intricate relation between mental content and the mind's external factors. These factors include at least the social and historical contexts in which persons live. Shared beliefs, languages, social practices, concerns, goals, and activities, all contribute to the kinds of minds persons have. For Davidson, the concept of 'pure objects of thought,' disconnected from these factors, makes no sense. Although minds have public and private aspects, neither aspect is primary to the other, as they are, in fact, mutually supportive. It is precisely on the basis of our engagements with other persons, and with the world of objects, that mental states are determined as the states they are. It is these engagements that provide the epistemological connections that give mental states meaning. According to Davidson, the very concept of raw experiential data is meaningless, as, to him, even the theoretical separation of data and conceptual scheme makes no sense (Davidson 1987), pp 159-172.

Davidson challenges the view that thoughts about oneself are more fundamental than thoughts about others. For him, this gets things the wrong way round. Rather than progressing from self-knowledge to other-knowledge, we must first acknowledge ourselves as members of a community of others, in a world in which there are other objects and other persons. Having thoughts involves having

[47] See, for example: (Davidson 1987), and (Malpas 1999).

beliefs, making judgements, holding propositions, and knowing the meaning of objective truth. It is only by conceiving of ourselves as persons among other persons that these things are possible. The recognition of our own viewpoint as one among other viewpoints permits us to judge, compare, make mistakes, change our minds, and therefore, to understand what it means to be 'objectively true' (Davidson 1995), pp 203-205. To formulate a proposition requires that one holds *other* propositions, which means holding beliefs about things in the world, other than just those things to which the proposition directly refers. Davidson points out the impossibility of understanding the belief that 'the sun is shining' without having other beliefs also. Such another belief could be the belief that the sun is in the sky above us, or that the sun appears dull when obscured by clouds, or that the sun disappears at night, and so on. Davidson holds that having these and similar beliefs presupposes a world of shared objects, within a common spatial and temporal frame, accompanied by a set of interpersonal relations (Davidson 1995), pp 211-217.

Davidson distinguishes between three different types of knowledge: self-knowledge, world-knowledge, and other-knowledge. All these types of knowledge concern the same reality, but from three different perspectives. Davidson uses the concept of 'triangulation' to explain these perspectives. We relate to others by means of a shared world of experience. In sharing this world, we share concepts, meanings, and language. These meanings correlate with objects and events in the world, thus forming a three way connection:

> Without this sharing of reactions to common stimuli, thought and speech would have no particular content - that is, no content at all. It takes two points of view to give a location to the cause of a thought, and thus to define its content. We may think of it as a form of triangulation: each of two people is reacting differentially to sensory stimuli streaming in from a certain direction (Davidson 1991), p 160.

These 'triangulations,' or common reference points, arise out of our common experiences. It is by being mutually experienced and understood that these reference points give our communications shared meanings. By sharing the world, we communicate with each other and understand each other's thoughts (Davidson 1991), pp 153-160. In explicating the externalist position, David shows how interpersonal communication and shared meaning is possible. But internalism notoriously faces problems in addressing such issues.

Work by Malpas on agency and spatiality also supports the externalist position. Central to agency is the implicit connection between the agent's mental content and the external world, or put another way, between the agent's intentional states and the world's objective causal ordering. The capacity for agency entails a grasp of this connection and of the fact that it is possible only in a spatially ordered world. Together, the above factors form the holistic structure which makes agency possible (Malpas 1999), pp 92-96.

The capacity for action requires an appreciation of the connection between agency and causation. Malpas notes that this understanding is not simply a matter of grasping what happens at the point of action; it also involves grasping the conditions under which agent-centred action is possible. These conditions distinguish agent-centred action from mere behaviour.[48] The conditions include the subjective state of the agent and the objective state of the external world. The first of these concerns the agent's capacity for intentional thought, while the second concerns the causal ordering in the objective world. For action to occur requires that an agent grasps both these conditions (Malpas 1999), pp 99-108.

Malpas argues that to act in any situation requires that agents appreciate 'causal bodily power,' that is, they are aware of the effect that a particular bodily action will

[48]For seminal work on the difference between actions and behaviour, see (Davidson 1980).

have. This involves having beliefs about oneself and one's capacities, about the causal connections between oneself and objects in the world, and about the causal connections between the objects within the world. These different beliefs cohere to form a framework of beliefs. To make sense, these beliefs require to cohere and connect with each other. One's belief about a particular action must cohere with similar beliefs about similar actions. The framework of beliefs is, then, one within which particular beliefs and actions make sense (Malpas 1999), pp 99-113.

Malpas notes that to have such a framework requires, not only a set of subjective states in an agent, but also a consistent causal ordering in an objective world. Unless there was some reliability of causal connectedness among objects, there could be no consistency of beliefs about objects. For example, to drink a glass of water involves my knowing how to lift my arm and how to make appropriate grasping and lifting movements in certain directions. It also involves knowing about the possibility of resistance to my actions. If my arms were immobilised by being tied in a sling, I need to be aware that my ability to act will be impeded. When I become aware of my limitations in this way, I become aware that my bodily capacities are related to objective facts in an objective world and not just to a set of subjective states in me (Malpas 1999), pp 114-124.

Malpas argues that an appreciation of spatiality is necessary to understanding the causal ordering in the world and how this ordering relates to action. Believing that actions at one time are causally connected to actions at another time means grasping those actions as distinct and temporally separated actions. But grasping this means also grasping the actions as located within a single spatio-temporal framework, in a single, objective world. To accept that one's actions at one time, are part of the same causal ordering at another time, requires that one appreciates them as both occurring within a single, spatio-temporal framework.

Just being an agent requires that one understands oneself as standing in a certain kind of causal and spatial relation to objects. This means having the relevant beliefs about those objects and about one's connection to those objects. If we did not have some kind of beliefs about objects and about our relation to objects, we would have difficulty initiating actions in relation to such objects. Grasping spatiality allows an agent to differentiate between the objective elements of the world and to understand how these elements are causally linked. An appreciation of spatiality also enables an agent to distinguish between her own intentional states and the states of the objective world. It is thus ultimately only through spatiality that agency is possible, and hence also, that thought is possible. The externalist conception of mind recognises this integral relation between spatiality and thought, but the internalist conception has difficulty in doing so (Malpas 1999), pp 114-137.

The above work elucidates ways in which internalism and externalism differ in their approach to mental content. Specifically, it shows that the externalist view can more satisfactorily account for the source of mental content than the internalist view. Because it acknowledges the preconditions of thought, externalism can also more satisfactorily account for psychological continuity. In linking the content of thought to the objects which first cause thought, externalism provides a more sound account of the content mental chains than a view which does not explicitly recognise this link.

3.4 Perception

In addition to philosophical argument, the externalist case is also supported by empirical research. In particular, research in the areas of Ethology,[49] Ecological

[49] *Ethology* is a biological approach to the study of animal behaviour, which is concerned with the animal's natural environment (Tortora and Becker 1978), p 777. For seminal work in ethology, see (von Uexkull 1934).This work is addressed in Section 3.5.

Psychology,[50] and Biology,[51] elucidates some of the ways in which mind and world are linked through perception. Perception is the interaction between a creature's perceptual apparatus and the world of objects, and the meaning given to that interaction. Perception, therefore, is the way in which humans and other creatures receive information from the 'outside' world. For humans, the information received through perception provides the material for mental content. Bodily interaction with objects, through the various sensory modes of hearing, smell, taste, touch, and vision provides the perceptual information from which thoughts are built. Without this sensory information, it would not be possible to think thoughts, not even *a priori* thoughts, as we would be unable to understand them. Perception, therefore, is fundamental to having a mind.

For humans, perception yields information concerning external objects, such as events, facts, persons, and other items (Honderich Ted 1995, p 652. This information forms the basis of our thoughts. Because thought and perception are so intimately linked, understanding perception is crucial to understanding the mind. To help to understand the link between perception and thought and to further support the case for externalism, perception is now examined from three different perspectives. These concern the perception of the immediate environment, the perception of objects in the environment, and the operation of visual perception.

3.5 Perception of the Environment

In its most general sense, perception is the means by which creatures receive information from the external world. It is perception, therefore, that guides and constrains a creature's interaction with its immediate environment. By perceiving

---

[50] *Ecological Psychology* is a psychology which recognises that the physical and biological features of the environment affect and influence mental life (Reed 1996), p 7. For foundational work in ecological psychology, see (Gibson 1977).This work is addressed in Section 3.6.

[51] *Biology* is generally concerned with the origin, history, structure and other features of the lives of plants and animals (Krebs 1981), p 84. Biology will explored in Section 3.7 specifically in relation to visual perception. See (Furst 1979), and (Atkinson and Braddick 1989).

their environment, creatures receive information concerning the factors that constrain their lives, such as the subjection to gravity, the tolerance of temperature, and the need for nutrition. Perception varies for each creature in accordance with its perceptual apparatus. For example, a starfish has no head or brain, but has rings of nerves around its mouth and thousands of neurosensory cells in the epidermis which are sensitive to touch, photoreception,[52] and chemoreception.[53] It also has a nerve cord in each of its five arms, and its main sensory organ is an 'eyespot' at the end of each arm (Curtis 1983), p 551. The 'information' given to the starfish through these various receptors guides its movement across the sea bottom, its search for food, and so on. The different perceptual responses of creatures can thus vary from the apparently simple,[54] like those of a starfish, to the very complex, like those of a human. The relation between creature, environment, and survival can be understood in terms of an interrelated framework. This framework comprises a creature's survival needs, its perceptual apparatus, its immediate environment, and its capacity for action. These factors concern the biological needs of organisms in an environment, and apply equally to humans and non-humans.

From an evolutionary perspective, these factors are taken to have influenced the emergence of minds. The need to deal with the complexities of multiple perceptual responses to the environment and the corresponding need to plan and execute action is thought to have influenced and guided the development of the higher cognitive faculties which characterise human minds. If minds first developed in this way, our understanding of the mind should include an understanding of the mind's biological underpinnings. Because it deals specifically

---

[52]'Photoreception' is the detection of light (Curtis 1983), p 1103.

[53]'Chemoreception' is sensory response to specific chemical stimulus, including smell and taste (Curtis 1983), p 1090.

[54]Although the perceptual capacities of some creatures may seem relatively simple when compared to humans, in their own right, they can still be very complex.

with creatures' perceptual responses to the environment, ethology provides an effective way of taking some steps towards gaining this appreciation.

While ethology has developed and become formalised in recent years, the foundational work in this area is understood to be that undertaken by von Uexkull.[55] Through simple studies and experiments, von Uexkull studied creatures in their natural habitats. His work produced some intriguing findings. He found that as a sensory being, a creature's perceptual response to its immediate environment is selective, and is part of a single system which composes the creature's world or *Umwelt.* The *Umwelt* comprises perceptual apparatus, motor capacities, survival needs, and perceptual response. This means that each creature or creature type responds to its environment according to its particular type of perceptual apparatus and its particular survival needs. By studying different types of creatures, von Uexkull learnt that due to the differences in perceptual apparatus and survival needs, each creature type perceives the environment differently to other creature types. A mollusc, for example, would have a completely different type of perceptual response to, say, a bird. Crucial to the present argument is von Uexkull's discovery that the set of environmental features which most concerns an animal[56] is that, and only that, to which it responds. This means that for humans, the kind of thoughts they can think is tied, by means of their perceptual responses, to the environments in which they are situated. Closer consideration of von Uexkull's work makes this link clearer.

As explained, to von Uexkull, the *Umwelt* or immediate environment comprises the set of environmental features relevant to, and consequently perceived by, any given creature. The features perceived by a creature relate to its life-activity and vary according to the needs of species and individuals. These features relate to

---

[55] Von Uexkull's ideas were later developed and formalised by others, such as Lorenz and Tinbergen.

[56] Humans are included as animals in this and other instances in this section.

the creature's nurture, maintenance, and development, and guide its selection of habitat. The *Umwelt* of each creature comprises its total perceptual response and hence its total world. Because each creature type responds differently, each creature type's 'world' is completely different to the 'world' of other creature types. What these worlds have in common, however, is that they are each constrained by the factors which permit them to exist. These factors include the distance between earth and sun, the basic elements found on earth, the climatic conditions, and so on. Together, these factors provide the totality of conditions under which the *Umwelten* exist. The totality of these conditions is the *Welt,* or the wider potential[57] universe (von Uexkull 1934), pp 5-6; (Ingold 1995), pp 62-66; (Eibl-Eibesfeldt 1975), p 6.

To explore the notion of the *Umwelt*, von Uexkull takes his reader on a stroll through 'unfamiliar worlds,' which are known to animals, but not to humans. We are asked to imagine each creature's world captured inside a soap-bubble, through which, looking outwards, we may see the world of the creature, our own world 'transformed:'

> When we ourselves then step into one of these bubbles, the familiar meadow is transformed. Many of its colorful features disappear, others no longer belong together but appear in new relationships. A new world comes into being. Through the bubble we see the world of the burrowing worm, of the butterfly, or of the field mouse; the world as it appears to the animals themselves, not as it appears to us. This we may call the *phenomenal world* or the *self-world* of the animal (von Uexkull 1934), p 5.

According to von Uexkull, the phenomenal world of the animal differentiates the animal from a mere mechanical object. Von Uexkull argues that unlike the physiologist, who regards a creature as no more than a machine, the biologist

[57] I say 'potential' here, because, it does not seem to me that it is possible for any creature to experience all the perceptual variations that are possible. In having one type of perceptual apparatus, one is effectively cutting-off the possibility of having another. Yet, because a certain set of physical conditions exist which permit all the creatures which *do* exist, to exist, we know that such a 'potential' universe, or *Welt must* exist.

appreciates a creature as a perceiving subject, a centre of a world, a being capable, albeit primitively, of phenomenally experiencing its own world.[58] Having different perceptual frameworks and limitations, different creatures, accordingly, have different experiential worlds (von Uexkull 1934), pp 5-9.

To bring out the difference between an animal and a mere machine, von Uexkull compares the perceptual response of a living organism with the mechanical response of a bell. When affected by temperature, or agents, such as acids, alkalis, or electric currents, the bell responds in the same way as any object which is constructed of similar metal. Living organisms, by contrast, respond to objects in accordance with the different cell groupings of which the organisms are composed. These cell groups form the organism's perceptual apparatus, or *perceptual tools,* and its operational capacities, or *effector tools.* Under the direction of a control centre or 'brain',[59] 'orderly collaboration' between the cell groups takes place. The receptor cells in the 'brain' receive and respond to stimulus, while the effector cells generate and control movement. The different cell groups perform their own functions, but by working in concert, they combine to produce representations of objects, which, for visually perceiving creatures could be 'the blueness of the sky,' or 'the greenness of the lawn.' Similar co-operation occurs in effector organs when motor impulses act on muscles to produce movement. In short, when a creature engages with objects in its environment, it responds perceptually, and generates its activity accordingly (von Uexkull 1934), pp 7-10. Von Uexkull sees a mutual relation between a creature's activity and its perceptual response, and the perceptual response and the creature's activity:

---

[58]Von Uexkull contrasts his view with that of some zoologists and physiologists who accord no more status than that of a machine to non-human creatures, and thus miss the awareness that many different worlds exist (von Uexkull 1934), pp 5-6.

[59]Quotes are inserted here around 'brain' as it is referred to in a very general sense and not specifically as the brain of a human or higher mammal.

> But since all the traits of an object are structurally interconnected, the traits given operational meaning must affect those bearing perceptual meaning through the object, and so change the object itself. This is best expressed briefly as: *The effector cue or meaning extinguishes the receptor cue or meaning* (von Uexkull 1934), the resulting arrangement or 'functional cycle' is a 'systematic whole,' an interrelated and mutually supportive set of factors, where an organism's perceptual response initiates its activity, and where its activity influences and defines its perceptual response (von Uexkull 1934), pp 10-11.

A creature's *Umwelt* comprises several such functional cycles. Von Uexkull notes that these cycles provides the first principle of '*Umwelt* theory,' namely, that all animals, whether complex or simple, have a reciprocal relation with, and therefore conform precisely to their own individual worlds. That is, there is a mutual relation between creature, survival, perception, environment and action - in short, an inherent interrelatedness between organism and environment, and environment and organism (von Uexkull 1934), pp 10-11.

Von Uexkull gives the example of the three functional cycles which compose the life-cycle of the female tick. The first cycle begins when the tick reaches maturity. Waiting in a suitable bush or similar place, the tick responds to the skin odour (butyric acid) of a passing mammal (first perceptual cue). Her response to the odour is to launch herself onto the unsuspecting prey (first effector cue). When alighted on the creature's surface, she seeks a hairless spot through which to burrow. The alighting action stifles the emission of butyric acid (second perceptual cue), causing the tick to 'run about' (second effector cue). On finding a warm spot (third perceptual cue), the tick proceeds to burrow and engorge herself with blood (third effector cue). Satisfied, she leaves her prey and returns to the ground. Here, she lays her eggs, and then dies. These three functional cycles comprise the tick's total life-cycle, and hence her total *Umwelt* (von Uexkull 1934), pp 7-11.

Von Uexkull claims that these cycles are not merely an exchange of forces, but are an exchange of relations between a living subject and its object. Although there is a multitude of possible responses from the tick, it 'chooses' the only three which are right for its needs:

> The whole rich world around the tick shrinks and changes into a scanty framework consisting, in essence, of three receptor cues and three effector cues - her *Umwelt.* But the very poverty of this world guarantees the unfailing certainty of her actions, and security is more important than wealth (von Uexkull 1934), p 12.

Von Uexkull caps his account by observing that a 'lucky coincidence' prevails between the scarcity of readily available tick food resources, and the tick's ability to survive for long periods without nutrition, up to eighteen years in fact (von Uexkull 1934), p 12.[60]

The simple life-cycle of the tick demonstrates the principles which apply to all creatures in their *Umwelten*. An interplay transpires between organism, needs, and environment, orchestrated by the creature's perceptual response and action. Von Uexkull relates how different types of visual responses result from different sets and types of visual apparatus. In each case, the apparatus supplies just what the organism needs for its survival. Together, needs, environment, response and action comprise the various creatures' *Umwelten*. In demonstrating the reciprocity between organism and environment, the functional cycles of von Uexkull's woodland creatures shed light on the biological underpinnings of human mind-formation and operation. Humans, like other creatures, also respond selectively to

---

[60]Von Uexkull deduces from the tick's unusual survival ability that creatures' experience of time is not uniform. At issue is the length of a creature's 'moment' - the amount of time in which the world 'stands still.' For humans this is 1/18 of a second. We can infer this by observing the amount of time required between frames to ensure that moving pictures appear as a continuous frame. Above this level (faster), the pictures would appear as a flickering confusion. Below this level, we could not detect movement (that is, with the naked eye).

Creatures do not experience time 'absolutely,' but in accordance with their relation to their *Umwelten*. This relation is the outcome of perceptual and effectual cues, which themselves are tied to the way creature and *Umwelt* relate (von Uexkull 1934), p 12. It can be seen that the process is mutually constructive and supportive, and hence is essentially holistic. Spatial response is also involved in this intricate relation, but its exploration goes beyond what I can reasonably undertake here.

their environments. The extremes of temperature we can tolerate are related to the survival capacities of our bodies and our vital organs. Similarly, the amount of food, water and sleep that we crave is related to our needs as biological organisms. Because these are the things which concern us, they are the things from which our thoughts are formed. Von Uexkull notes that because our response to the world is selective, we can often have a 'search image' in mind when looking for objects. In some instances, this 'search image' is powerful enough to 'block' us from seeing actual images. Von Uexkull gives the example of a clay pitcher which he frequently saw in a particular spot on a friend's dining table. One day, the pitcher was broken and replaced by a glass jug. Not expecting the unfamiliar image, von Uexkull was completely unable to see the replaced item until it was pointed out to him (von Uexkull 1934), p 62. This simple example demonstrates how our needs influence and often determine what we perceive, and therefore, what we think.

Von Uexkull gives many examples of creatures' different phenomenal worlds. The phenomenal world of one creature differs from that of creatures differently configured. In each case, reciprocity between creature, perceptual response, and environment is demonstrated. A single scene, for example, is perceived differently by a man, a dog, and a fly. Of course, we cannot *know* for certain that the phenomenal content of another creature's perceptual experience is like the pictures offered by von Uexkull, but this is beside the point. What we can appreciate is that a creature's perceptual world is the outcome of an interrelated set of factors, and that correspondingly, so is our own (von Uexkull 1934), pp 20-72.

The selectivity of *Umwelt* response is brought out by von Uexkull's example of a single oak tree. To the creatures who live there, it is home, protection and sometimes, nourishment. Each creature who lives in the tree chooses a different

part, according to its needs. The ant chooses the crevices of the gnarled bark, the boring beetle the bark's underside, the fox the roots, and the owl the branches. In each case, the creature's response to the tree relates specifically to the creature's own needs and concerns, and furnishes each creature with its own unique world (See Figures 1-4).

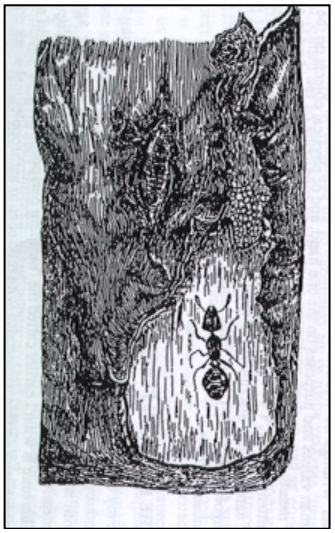

Figure 1 Ant

Figure 2 Boring Beetle

Figure 3 Fox

Figure 4 Owl

Ant, Bark-boring Beetle, Fox, Owl and Oak Tree. Illustrations by G. Kriszat.
Reproduced from (von Uexkull 1934) , pp 76-79.

While there is only one tree, each creature responds to different parts or aspects of the tree. Each creature selects its habitat from the part of the tree which best supplies its comfort, nutrition, protection and suitable vantage point for watching predators. The creature's perceptual response delimits this complex arrangement in each case. Most likely, each creature is unaware of the perceptual responses of

the other creatures, and may even be unaware that the other creatures exist (von Uexkull 1934), pp 74-76.

Just as the tree appears differently to the various creatures who live there, the tree also appears differently to humans (See Figures 5-6).

Figure 5 Girl Figure 6 Forester
Girl, Forester and Oak Tree. Illustrations by G. Kriszat. Reproduced from (von Uexkull 1934) , pp 74-75.

To the little girl, the tree is a frightening demon. To the forester, it is just the next tree to be felled. He is oblivious to the frowning, forbidding face. These examples demonstrate how a single object can elicit a variety of different perceptual responses and how these responses are related to the different perceptual apparatus and the different concerns of the creatures involved. As a result of their different concerns, each creature perceives the one, single object, differently. We can never know what these different perceptions are actually like. Thomas Nagel was aware of this when he argued that we could never understand what life was like for a bat (Nagel 1979), pp 169.[61] We can speculate though, that the bat

[61]See Nagel's seminal article: (Nagel 1979). Nagel argues that the different needs and modes of perception of other creatures, such as bats, ensure that their phenomenological experiences are different from ours: 'Bats, although more closely related to us than those other species, nevertheless present a range of activity and a sensory apparatus so different from ours that the problem I want to pose is exceptionally vivid . . . . bat sonar, though clearly a form of perception, is not similar in its operation to any sense that we possess, and there is no reason to suppose that it is subjectively like anything we can experience or imagine' (Nagel 1979), p 168. Although we cannot imagine precisely *what* it is like, we can be aware of it as a field of activity and experience which is necessarily different to our own.

occupies, and responds to, a set of environmental features - an *Umwelt* - quite unlike our own. Similarly, the world of a bee, a snail, or other creature, cannot be a world for us.

For von Uexkull, the oak tree and its diverse inhabitants provide a mirror to the rest of nature, in which multitudinous possibilities are present:

> What we have found on a small scale in the oak tree is enacted on the life tree of nature in vast dimensions (von Uexkull 1934), p 76.

The 'dimensions' perceived by humans differ in accordance with their particular needs and concerns. Von Uexkull gives the example of different types of scientists, each of which sees the same world through different eyes. Because of their different concerns, they each 'see' the world differently to the others. For the astronomer, the world is one small object among millions of other objects in space; for the deep-sea researcher, the seas of the world contain nature's abundance; for the nuclear physicist, the world comprises minute particles. The colours which these scientists see are seen differently by the physiologist under his microscope. And the sounds they all hear are heard differently by the tuned ear of the musician. Although each of these 'worlds' are different, they are all underpinned by the same natural laws, and thus, at a fundamental level, are constrained by one nature (von Uexkull 1934), pp 76-80.

Von Uexkull's account of the selective responses of humans and other creatures to the one, single world, draws attention to the mind's biological underpinnings. In particular, his work treats organisms as always in systematic relation to their environments. This means that their individual responses have to be understood in relation to the environmental context that elicits them. Individual responses cannot be understood independently of environmental circumstances. If we were to extend this to mental states, it would suggest that they cannot be viewed independently, not only of their environments, but also of the bodily characteristics to which they are related, and by means of which creatures hold themselves in

contact with their world. This means that the kinds of thoughts we think, indeed the kind of thoughts that we *can* think, are, at least in part, a direct product of the kind of world in which we live, and of the kind of interaction we have with that world. Another way of putting this is to say that the world is inside our minds, rather than our minds being inside the world. The psychological continuity criterion fails to recognise the connection between our minds and the world. Indeed, in confining its account of the mind to an account of 'overlapping chains,' the psychological continuity criterion fails to acknowledge that the world not only constrains and influences our thoughts, but also, our capacity to think at all. The importance of environmental influences on our perceptual abilities is further taken up by J. J. Gibson.

3.6 Perception and Affordance

The relevance of environmental features to mind formation is also emphasised by ecological psychology. In particular, Gibson's work on *affordances* relates environmental features to creaturely activity and to the objects involved in that activity. An affordance is any item in the environment which relates, either directly or indirectly, to activity. Affordances contribute to a creature's development and maintenance, and include natural and manufactured objects. They apply to all creatures including humans, and generate a reciprocal relation between creature and environment.

Because they are intimately tied to perception, affordances are relevant to human experience, and hence also to mind formation. To understand how affordances influence and constrain minds, we need to appreciate how they are involved in a reciprocal arrangement between any creature and that creature's basic biological needs. While humans are thinking creatures, they are also biological creatures, and are, therefore, involved in this arrangement also.

According to Gibson, an affordance is:

> a specific combination of the properties of its substance and its surface taken with reference to an animal . . . . a combination of physical properties of the environment that is uniquely suited to a given animal - to his nutritive system or his action system or his locomotor system (Gibson 1977), p 67; p 79.

Aspects of the environment which are useful to a creature's maintenance constitute its affordances. Items valuable to a creature's nutritive system are determined by its being a herbivore or carnivore; those valuable to its action system concern whether it has claws or hands; and finally, those valuable to its locomotor system relate to its type of legs and feet (Gibson 1977), p 79. Affordances range from simple to complex, examples being 'features of the terrain, shelters, water, fire, objects, tools, other animals and human displays.' The conditions under which these items are perceived are also affordances, an example being the presence of sufficient light for vision to be possible (Gibson 1977), p 67.

Gross environmental features are affordances because they determine the type of creature able to live in a given environment. For example, a land surface, solid and flat, better supports large or heavy animals than one swampy and sunken. Mobility and upright posture is sustainable, legs can walk and stand, conditions are thus conducive to quadrupeds and bipeds. An affordance could also be a seat. It is the means of 'sitting down' because, having the right solidity, shape, surface, and height, it can support the required part of the human (or other) frame. Affordances are also objects that can be climbed on, crawled under, moved around and so on, or otherwise encountered by creatures in the course of their pursuits. More generally, affordances are the tools, shelter, food, and drink provided by the environment to creatures in accordance with creatures' manipulative skills, shelter needs, digestive systems, and so on. Together, these

features compose a *niche,* that is, the set of affordances selected and utilised by a creature from its environment, according to its own particular set of needs. Gibson makes the point that although a niche implies reciprocity between creature and selected features of the environment, it also implies the presence of an 'unlimited environment,' from which these features can be selected (Gibson 1977), p 68-69. An affordance is thus more than an arbitrary subjective response to one's immediate surroundings, but is also an objectively identifiable construction, in an objectively identifiable world. For humans, the set of environmental features which form their niches are those to which they respond, and thus, are those to which their minds become 'tuned.' Were humans to be differently constructed, their niches would be different, and consequently, their thoughts and identities would be different also.

For niches to exist requires that amenable geographical, chemical, and climatic conditions also exist. Physical features such as caves, water-holes, swamps, rivers, river-banks, plains and so on must be available. Materials for habitat construction (human and other), must also be obtainable. Compatibility between creature and niche is necessary, as each 'fits' the other. Unless this occurs, creatures could not survive, as it is the elements comprising niches which allow them to do so. Human capacity to alter the environment is extensive, but not unlimited. While modification of the earth to better serve human needs has been widespread and cumulative, it is, like aspects of all other life forms on the earth, ultimately constrained by the nature and composition of the earth itself (Gibson 1977), pp 69-70. Similarly, the earth (and, by extrapolation, the wider universe and the scientific laws within which the earth exists and operates) in constraining what is available for our niches to exist, also constrains what is available for us to think about.

Elements on and in the earth contribute variously to the life forms sustained. Elements and combinations of elements are fluid or solid, and roles vary accordingly. For example, oxygen sustains breathing, transparent air sustains vision, and water sustains drinking. More solid substances can be eaten or provide materials for tools or habitats. Materials are subject to manipulation to varying degrees, and provide life-support or life enhancement. Gravity anchors some creatures to the earth, while permitting others to fly from it without being catapulted into space. Earth provides the requisite conditions for lying, and for locomotion, such as walking or running. Due to 'cluttering,' this latter feature is not ubiquitous, as some surfaces, such as swamps or cliffs, require different methods of negotiation. For many creatures, vision prevents collision, and thus aids locomotion. Because vision is limited to the perceiver's visual field, vision also implies concealment, meaning objects can be hidden as well as seen. This could not occur unless some surfaces were opaque. Gibson points out that the very possibility of concealment contributes to the idea of enclosure and privacy (Gibson 1977), pp 71-74. These features of the earth are significant to humans, as they relate to the structure of our bodies, to our capacities to negotiate ourselves through the world and to our biological needs. Like any creature, we do not respond equally to all of these features, but only to those which concern us, and which relate to our bodily structure, survival needs, and perceptual apparatus. Thus, although these features of the earth are concealed in a superficial analysis of thought, they provide crucial material for our thoughts and are part of what makes our thoughts possible.

Object manipulation occurs in accordance with the availability and usefulness of objects, combined with the manipulative capacities of creatures, whether human or others, such as chimpanzees. Objects of appropriate size, shape and solidity provide the means of performing tasks, such as levering, throwing, spearing, cutting, tracing, and so on. The affordances which most impact on any creature

are other members of its own species, as it is these which provide the most direct influence on behaviour, through interaction of varying kinds, such as nurturing, or sexual activity. Affordances thus include all aspects of a creature's environment which, in one way or another, contribute significantly to the creature's survival (Gibson 1977), pp 74-76.

In summary, the notion of affordances draws attention to the point that objects, both natural and manufactured, elicit differential responses from different creatures in the 'same' world. An object, or feature of the terrain, is perceived by a creature in accordance with the creature's needs and concerns. This indicates that perception is neither autonomous nor arbitrary. What a creature perceives is the outcome of its perceptual apparatus, its survival needs, and the particular selection of items in the environment to which it responds. This is equally true for humans and non-humans. Human sensory modalities operate in conjunction with human needs, concerns, and abilities. We do not readily (or even at all) pick up from the environment all that is there. We note and respond to what is most important or significant to our 'niche.' We are blind to the features which interest our cat or our goldfish. But we are concerned that the supermarket stocks our food, that the petrol-station is open, and that the road is not flooded. Because these items concern us, they 'register' with us, we respond to them, and they become part our mental repertoire. These items are the objects from which our thoughts arise, and consequently, they are the very stuff from which our minds are made. If we are to comprehensively understand our minds, and especially, the *contents* of our minds, we need to understand how our minds relate, not only to our environments, but also to the particular features of the environment that interest and concern us. These points are important when accounting for personal identity, for if the external world is crucial to the constitution of minds and the contents of thoughts, then it is crucial to the formation of personal identity also. The psychological continuity criterion is remiss in this regard, as when accounting

for personal identity, it fails to take any of these factors into account, and thus fails to include reference to many of the most crucial features which personal identity involves.

3.7 Visual Perception

The specific modalities by which humans perceive the world are also significant to mind formation, and correspondingly, to personal identity. Perceptual information is received from the immediate environment in the five modalities of sound, smell, taste, touch, and vision. While each modality provides a different type of information, there are common principles involved. Each modality involves an interaction between the perceiver and the world by means of the appropriate perceptual apparatus. Cognitive systems are involved in the interpretation and understanding of this interaction. Visual perception is, in most circumstances, a primary source of perceptual information for humans and other creatures. The process of visual perception in humans provides further evidence of ways in which our perceptual interaction with the world of objects determines and constrains the nature of our thoughts, and hence, the contents of our minds. Perceptual input for sighted persons is considerably different to perceptual input for non-sighted persons. We thus could not expect the structure or content of thoughts for both to be the same, (although this does not mean that there could not on occasions be some thoughts which might be very similar, such as 'I am cold' and so on). Understanding the visual process and how it relates to thought gives insight into the integral connection between perception and thought, and correspondingly, between perception and personal identity.

Studies show that visual perception is not a discrete act, but involves a complex array of factors. One key factor is object recognition. To recognise objects requires an interaction between the perceiver's visual systems, and the perceiver's stored knowledge. This knowledge results from practical engagement,

either directly or indirectly, with the objects concerned. It is this engagement, or connection with engagement, which provides meaning to the perceiver. For example, knowing that an object is a chair involves more than the technical operation of one's visual apparatus. It also involves knowing the purpose of the chair. This, in turn, means 'knowing' about the difference between sitting and standing, having legs, having knee joints, and so on. I come to know about these things by incremental activity, beginning in infancy. If I did not know about these things, I would not know how to identify the object as a chair. I could mistake it for a coffin. If I was born deformed and unable to use a chair myself, I could learn about the chair from watching others.

Perceivers do not recognise objects in vacuums, but from accessing stored knowledge, either about, or related to, the object concerned. Without this knowledge, objects would have no meaning, and thus, would be unrecognisable. Interacting with objects is thus crucial to our ability to interpret our perceptions and understand our thoughts. This need to interact with objects is not acknowledged by the psychological continuity criterion, as it takes our ability to understand objects for granted.

Another key factor in the visual process is the human visual system's developmental nature. Visual systems are underdeveloped at birth, and would remain so unless appropriately stimulated. This stimulation occurs when an infant first begins to respond visually to the world of objects. The infant's interaction provides the necessary conditions for the physical structures of the eye and brain to mature, *at the same time as* providing the necessary conditions for the development of semantic networks. Thus, the maturing of vision and the development of the knowledge of objects occur in tandem. This point is supported by the fact that persons who first acquire vision as adults are unable to

immediately recognise objects.[62] The successful development of vision requires that interaction with the world of objects takes place.

Visual systems comprise retinal and brain structures. The visual process is initiated by the reception and regulation of light by the iris. Receptor cells in the retina then transform this light into electrical energy. Next, an image is focused by the lens onto the retina. Vision is clearest at the central part of the retina, the *fovea*, where the receptive cells are most densely packed. Reflexes in the brain control the eye's movement. When the eye surveys a scene, its blinking movement means that the brain receives information only intermittently. Yet, the brain is able to integrate these incomplete signals into a coherent picture. This process occurs due to several complex factors (Furst 1979), pp 36-39. When we understand these factors, we realise that perception and thought occurs in the first instance, only in virtue of our interactions with the world of objects.

Like the brain, of which it is part, the retina is neurologically complex. It has several layers of orderly, functioning nerve cells, an important task of which is to perceive contrast, thereby detecting visual edges. The visual cortex is connected to the back of the eyes by the fibres of the optic track. The central section of fibres from each eye cross over en route, so that information from the right half of each eye reaches the left hemisphere, and information from the left half of each eye reaches the right hemisphere. At least 80 percent of the cortex nerve cells respond to stimulations from either eye. Each retina half relates to the part of the visual cortex to which it is connected.
Thus, when points on the retina are activated, so also are the corresponding points on the cortex. This can lead to the mistaken conclusion that visual perception consists of an 'homunculus' or tiny man watching a film inside the brain Furst 1979), pp 37-40.

---

[62]This claim is more fully explicated shortly.

Fibres from the optic nerve also connect to the *optic tectum,* situated at the back of the brainstem. For creatures without a visual cortex, such as birds and amphibians, the optic tectum forms their total visual brain. For humans and other mammals, the optic tectum informs of peripheral objects, while the visual cortex provides more detail about centrally placed objects. Vision involves the division of a scene into its constituent parts. Recognition take place, allowing a scene to be *seen.* The ability to 'see' an object is not just a matter of the operation of a set of visual apparatus. 'Seeing' alone is insufficient for recognition, as becomes apparent when we realise that some scenes can be interpreted in more than one way. This would not be possible if visual apparatus was all that was involved.

The picture below makes this point clearer (See Figure 7). We interpret the picture either as a vase, or as two symmetrical faces. Neither of these interpretations would be possible unless we already had some basic knowledge of similar objects. Such knowledge could only come from having interacted with these objects. It could not have come 'out of nowhere.' This example shows how factors apart from, and external to, the visual process itself are involved in the act of recognition.

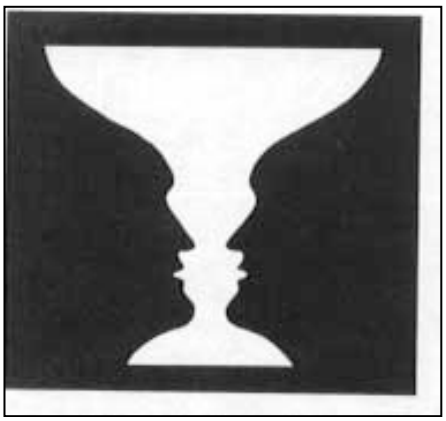

Figure 7 Vase or Two Faces
Reproduced from Charles First, *Origins of the Mind* 1979, Prentice Hall, New Jersey, p 41. The picture can be recognised in two ways. This shows that the brain segments the picture in accordance with already acquired semantic content. It does not interpret the scene ex nihilo (Furst 1979), p 41.

Recognition is essential, as otherwise no sense could be made of visual presentations. It was first discovered in 1932 that former cataract sufferers cannot, after cataract removal, immediately understand their visual perceptions. They require time to become familiar with the intricacies of sight before they can identify even simple objects, such as basic shapes. They can only achieve this familiarity by interacting with objects, and learning to match the unfamiliar visual input with the already familiar tactile input (Furst 1979), pp 39-42. It is apparent that the ability to successfully recognise objects is a crucial aspect of effective vision. Computer experiments show that this successful recognition is more than simple template matching,[63] such as occurs in banks when customer numbers are read from cheques or credit cards. Because it lacks semantic awareness, a computer could fail to recognise even a simple object such as a comb (See Figure 8).

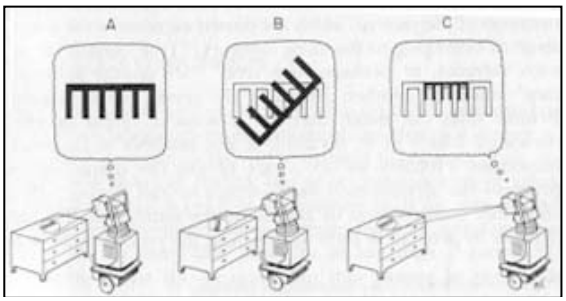

Figure 8 Computer with Comb

Reproduced from Charles First, *Origins of the Mind* 1979, Prentice Hall, New Jersey, p 45. The comb is unfamiliar to the computer because it is oriented differently to the computer's program, and thus recognition of the comb by the computer fails (Furst 1979) , p 45.

Factors such as orientation could cause the comb to be presented to the computer differently to the format of the computer's template. Recognition by the computer may then fail. Human vision, however, is more complex than this. A basic categorisation process occurs prior to the specific recognition of an object. This involves matching input with a basic schema, such as 'dog', 'tree', or 'person.' Considerable latitude is required, as individual objects in the world differ and are unlikely to match exactly with an idealised form. The matching process works only because it draws on stored memories of similar objects, and accompanying

[63]A template is an idealised form or pattern guide to which objects are formed or matched.

semantic content. Object recognition is not just a matter of perceiving a particular form, but of attributing meaning to that form. Tests on animals, in which individual neurones are monitored with tiny electrodes, at the same time at which objects are presented visually, demonstrate that brain cells respond to specific types of stimuli, and also to specific parts of the retina. Due to the experiences which occur during visual development, brain parts are 'tuned' to respond in particular ways. For example, due to incremental activity coded in the brain, frogs, who possess only the visual tectum, respond to moving flies, but are incapable of responding to stationary ones. As animals become more evolved, their ability to vary their responses increases. For humans, development includes the capacity to code visual or other inputs with comparatively abstract schemas. The recognition of alphabet letters, even when in heavily ornamented, disproportionate, or unfamiliar form, is one example (See Figure 9) (Furst 1979), pp 43-53.

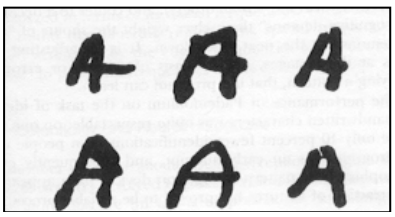

Figure 9 Alphabet

Reproduced from Charles First, *Origins of the Mind* 1979, Prentice Hall, New Jersey, p 53. Imprecise letters are recognised because the mind refers to what it already knows from previous interactions with similar objects (Furst 1979) , p 53.

Object recognition requires the ability to distinguish between shades of light and dark, even when the differences involved are minute. Groups of neurones respond to light falling in the particular receptive field to which the neurones are tuned. Edges and lines are detected at the visual cortex, while receptive fields at the secondary or prestriate zone respond to these edges and lines also, in addition to the larger region in which these edges and lines are situated. The role of recognition and interpretation is illustrated by the initial inability to detect figures in camouflage, but following successful detection, the subsequent inability to disregard them (See Figure 10).

Figure 10 Camouflage

Reproduced from Charles First, *Origins of the Mind* 1979, Prentice Hall, New Jersey, p 44. This picture demonstrates how recognition of familiar items requires an active process of interpretation. Can you recognise the figure in the picture? It is usually difficult at first, but once seen, cannot be missed. Try to find the figure before reading the footnote (Furst 1979), p 44.[64]

The spatial frequency of image components is also important to recognition. Relevant here is the *spatial-frequency spectrum,* by which a pattern's decompositional features of brightness and darkness are measured in cycles per second. The reproduction, in block form, of well-known painting demonstrates this aspect of recognition (See Figure 11).

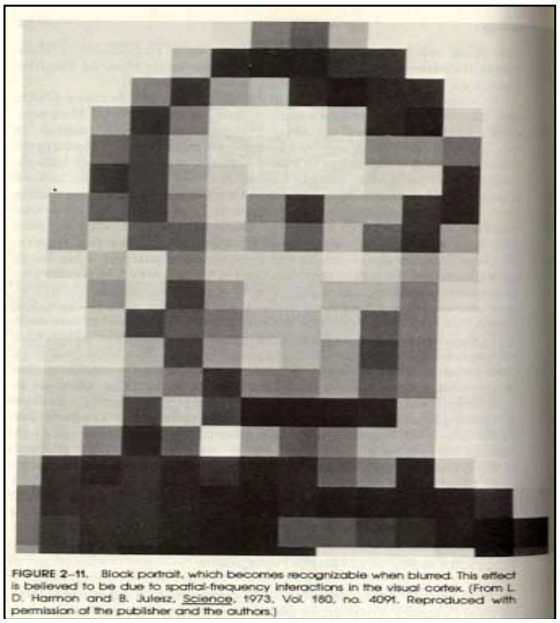

Figure 11 Block Face
Reproduced from Charles First, *Origins of the Mind* 1979, Prentice Hall, New Jersey, p 56. The picture becomes recognisable on squinting, but only because we have stored memories from having previously seen and understood the image (Furst 1979), p 56.

[64]The hidden figure is a man carrying and pointing a rifle.

Brightness values are averaged in blocks throughout the painting. Although this masks the intended image, blurring one's vision by squinting allows the unnaturally emphasised edges to be removed, thereby permitting the face to be revealed (Furst 1979), pp 55-56. Unless the face was already familiar to us we would not recognise it. Other strange shapes, such as the one below, reveal the role played by memory in process of visual recognition (See Figure 12).

Figure 12 Hidden Face

Concentrate on the four dots in the middle of the picture for about 30 seconds. Then close your eyes and tilt your head back. Keep them closed. You will see a circle of light. Continue looking at the circle. What do you see? The disparate segments of the image only cohere into a recognisable face because it is one with which we are familiar, having seen the image, or something like it, previously.[65]

The above observations of the visual process reveal that the achievement of vision is not a discrete or simple process. Vision occurs due to the interaction between the perceiver's visual apparatus, the perceiver's mind, and the perceived object. At least four pertinent features are evident in this interaction. First, vision entails object recognition. Second, object recognition draws on the perceiver's stored knowledge. Third, this knowledge is attained from direct or indirect interaction with objects 'outside' the mind. Fourth, the *type* of vision which occurs is directly tied to the particular kind of visual system that we have. For example, certain aspects of vision, such as the ability to detect lines and edges, are only possible due to the particular physical constitution and construction of our visual

[65] I have been unable to obtain a reference for this picture. It came to me indirectly from my daughter's school-teacher, who has been unable to discover its source.

apparatus. The combination of all the above elements is required for successful vision to occur. And because successful vision is dependent on the combination of these various elements, the thoughts which arise from vision are dependent on these elements also. Thus, neither our perceptions nor our thoughts can be adequately understood unless we refer to the world of objects from which those perceptions and thoughts first arise. Correspondingly, neither can personal identity be fully understood unless we refer to the world of objects which underpins our perceptions and thoughts, and makes them possible. But in maintaining an internalist view of the mind, the psychological continuity criterion is unable to account for the connection between thought and the objects in the world, and thus, its account of personal identity is disconnected from the world of objects in which we live. The dynamic nature of the visual process becomes more apparent when developing visual systems are considered.

The progressive nature of pattern recognition is evident from the study of infant visual development.[66] Pattern recognition is not immediate, but is the outcome of an extended process, which involves an interaction between the infant and the world. Infants' responses to visual stimulation situations indicate their biological readiness to investigate their surroundings. Their perceptual abilities arise from a combination of rudimentary abilities present at birth, and subsequent learning, necessary to which is interaction with the world (Slater 1989), pp 43-44. Vision achievement is particularly relevant to normal cognitive and social development, as it is one of the early key ways in which information is extracted from the environment. Investigation of this process shows that rather than being a static function, operating uniformly between birth and death, vision begins in a primitive form in the newborn, and (normally) progresses through various stages to a

[66]This discussion of infant visual development is not intended to be comprehensive. The process is complex and detailed, and cannot be fully investigated here. The ideas addressed are meant to give an example only of some of the principles involved. For an extensive coverage of the main issues, see (Slater and Bremner 1989).

matured form in the adult. The achievement of acuity is one example of this development (Atkinson and Braddick 1989), p 7.

Developed in the early stages of vision development, acuity is the ability to refine spatial detail. Observation of retinal images in newborns reveals that infants have the capacity to perceive nearby objects, such as their mother's face, but that they have an underdeveloped sense of acuity. While optical structures function adequately in newborns, neural structures are underdeveloped, generating faulty *accommodation* responses. 'Accommodation' refers to the eye's ability to bring objects at different distances into sharp focus. Optical structures are responsible for receiving information, while neural structures are involved in processing this information. Although the information is correctly received, immature neural structures inhibit its processing, resulting in blurred vision (Atkinson and Braddick 1989), p 9.

Immaturity in the structure of the eye's cone receptors is also involved in the lack of visual refinement. In adults, these receptors are concentrated in the central area of the eye (fovea), and are densely packed. They are responsible for visual clarity. In infants, they are less concentrated, and therefore less centralised, resulting in impaired visual resolution. Visual exposure stimulates the receptors, initiating repeated saccadic (rapid flicks) eye movements by one month-olds. These movements appear to anticipate the ultimate centralised location of the cone receptors, and are seen as failed efforts to fixate visual targets in the central position, where they will be located once the cone receptors are in their final positions. Practice in target location over several months contributes to the continuing development and co-ordinated functioning between both optical and neural components, resulting in the infant's visual improvement, and the ultimate achievement of acuity (Atkinson and Braddick 1989), p 21-22. Although infant visual development is considerably more complex than this, the process outlined

above is one example of how engagement in the visual process itself is integral to vision's ultimate successful development. As can be seen, this process is dynamic, and entirely dependent on certain interactions with the world of objects for its successful outcome. Without such interaction, vision could not occur, and consequently, neither could the thoughts that arise from vision.
The visual processes which underpin these thoughts are taken for granted by the psychological continuity criterion, which pays no attention to the preconditions of thought, and to the fact that thoughts do not arise out of nowhere, but arise only as the result of the particular interactions of particular bodies, in particular surroundings, at particular times. This neglect of the external factors which are involved in the preconditions of thought results in an internalist view of the mind, and consequently, in an interpretation of personal identity which is inadequate and misleading.

The importance of engaging with objects to successful visual development is borne out by studies of some eye disorders. *Amblyopia,* a functional loss of visual performance, and *strabismus* (squint), the permanent misalignment of the two eyes, are examples.[67] A common form of human amblyopia, *deprivation amblyopia,* results from cataracts covering one eye and clouding vision. Because the damaged eye is less involved in vision, the nerve fibres which convey signals from the eye to the neurones in the visual cortex are less stimulated than those of the functioning eye. As a result, the damaged eye becomes functionally disconnected from the brain. Animal experiments, in which one eye is covered or blurred during the early phases of visual development, demonstrate this point. These experiments show that most of the immobilised eye's connections to the visual cortex become destroyed. This indicates that there is a direct link between an eye's stimulation and the eye's successful connection to the brain. Eyes, which through cataracts or other impairments, fail to receive this stimulation, are unable

[67]Between 2 and 8% of children develop these eye disorders (Atkinson and Braddick 1989), p 24.

to establish and support this connection, and thus become functionally inert (Atkinson and Braddick 1989), p 24. This example brings home the point that the visual process itself should not be taken for granted, as it is dependent on the occurrence of certain interactions between a person's visual apparatus and the world of objects.

If these interactions did not occur, vision would not occur either. Vision is not something which is fully developed at birth, but is a process which requires interactions with objects for certain components of the visual apparatus to mature. We can see from this that our perceptual input, and correspondingly our thoughts, are not discrete from our interactions with the world in which we live, but are completely dependent on those interactions.

Vision is also affected in cases of strabismus, where the eyes are misaligned. In animal studies where eyes become misaligned artificially, it is found that eye input to the visual cortex occurs, but that the cortex is abnormal. Neurones are missing, which would normally combine the input from both eyes. This deficiency is thought to result from the eyes' misalignment. Because the eyes are differently situated to each other, the stimulus between the eyes is unequal, resulting in the absence of correlated activity between the eyes. The regular occurrence of correlated activity between both eyes is thought to be required for successful binocular vision to develop. This correlated activity may also be linked to other important brain developments associated with successful vision (Atkinson and Braddick 1989), p 26. However, as this activity is unable to occur, correct visual development cannot occur either. Thus, perceptual input and thought are reduced accordingly. That thought can be limited in this way is missed by the psychological continuity criterion, as it does not recognise the various interactions between perceptual apparatus and world which are essential to normal perceptual input, and consequently, that certain thoughts are entirely conditional on such input.

These findings concerning both amblyopia and strabismus indicate that the successful development of visual perception is directly tied to the amount and type of stimulus received by the eyes during the developmental process. This stimulus only occurs as a result of interaction with objects which exist in the world 'outside' the mind, and thus, this stimulus an important component of visual perception, and consequently, of the thoughts which result from such perception.

In summary, the investigation of visual perception indicates that vision is not a simple or discrete process, but is a complex and interactive one. Parts within sets of visual apparatus are complex, and operate by mutual interaction. Successful vision results from the interaction between perceptual apparatus, stored knowledge, and objects. This combination of factors is required for object recognition to occur. Vision is immature at birth, requiring specific developmental processes for vision to mature successfully. In tandem with these developmental processes, infants learn about their surroundings, and become acquainted with the world of objects. For both adults and infants, vision is the outcome of specific interaction with objects, without which, vision could neither be achieved, nor developed. Because object interaction is crucial to successful vision, object interaction is also crucial to the mental content which vision provides. This means that the contents of our thoughts are directly related to the kinds of interactions we have with the world of objects. These interactions are guided and constrained by our perceptual apparatus, as it is only in virtue of that perceptual apparatus and the particular way it operates, that we are able to receive information from the world, and thus, are able to think thoughts at all.[68] Although the focus here has been on visual perception, similar findings are likely to apply to the remaining perceptual modalities, as they too form part of the single system of environment,

[68]Of course, an objector can always say: 'How do we know that our minds would be different if our visual process was different?' The point I want to make here is that if we take this attitude to everything in the mind, we will never learn anything about the process of thought and how it relates to the world. In my view, we need to focus more on what actually happens, rather than on what might happen, if we are to learn anything constructive about personal identity.

perceptual response, motor capacities, and survival needs which form the *Umwelten* of humans and other creatures.

3.8 Summary

I have argued in this chapter, that by holding to 'any cause,' the psychological continuity criterion maintains an internalist view of the mind. On this view, the mind's external factors are not explicitly specified in the account of mental content. All that matters is the presence of the ‘right kind’ of causal connectedness between items of mental content, regardless of how that mental content came about in the first place. This means that the actions, events, items, places, relations, and so on, which furnish the material for thought, are not included in the account of thought. The problem with this view is that it assumes that mental content is, in some way, just 'given,' and is in no way tied to factors such as bodies, survival needs, or environments. Even were these factors to be considerably different to what they are, mental content, on this view, would remain unaltered. But in holding to this view of the mind, the psychological continuity criterion does not adequately account for the preconditions of thought, and is thus incomplete in its account of the mind, and correspondingly, in its account of personal identity.

I have also argued that in contrast to the internalist view, the externalist view acknowledges the objects which first cause thought, and is, therefore, more comprehensive in its account of the mind, the contents of minds, and the conditions under which minds are possible. By including reference to the actions, events, items, places, relations, and so on, which exist 'outside' the mind, the externalist view is more thorough in its account of the mind than a view which does not explicitly refer to these things. I have argued that if there were no objects 'outside' the mind, there would be no material for mental content and minds would not be possible. Due to the intimate link between minds and the external world, it

is false and misleading to conceive of persons apart from the worlds in which they are situated, and apart from their particular form of perceptual interactions they have with that world. To support the externalist case, philosophical argument and empirical research from a variety of sources has been considered.

The work of Davidson and Malpas shows how the intelligibility of thought requires that certain background conditions exist within which thoughts are possible. These conditions include a framework within which thoughts cohere with other thoughts, and within which relations and communications with other persons are conceivable. Because thought relates to action, reference to the conditions under which action is possible is also required. Key among these is an appreciation of the difference between the subjective states in an agent, and the objective causal ordering in the world, as it is these things which permit agent-centred action to occur. The requirements specified by Davidson and Malpas are met through our engagement and interaction with the objects, persons, and places which exist outside the mind, as it is these engagements and interactions that provide the basic material for mental content and make minds possible.

The empirical research focused on perception. Studies were considered, which indicated that having a mind is related to our status as biological creatures. Our perceptual response to our immediate environment and to the objects we use, is, like that of other creatures, related to our survival and other needs. Because our thoughts are intimately connected to our perceptions, our thoughts are also linked to the conditions and objects to which these needs relate. Similarly, the consideration of visual perception showed that interaction with objects is a crucial aspect vision, and that accordingly, object interaction and the visual process are crucial to the thoughts initiated by vision.

The above philosophical arguments, along with the examination of environmental perception, object perception, and visual perception, supports the case that external factors are crucially implicated in mental content. When we realise that thoughts initially arise from objects 'outside' the mind, we become aware that these objects are a necessary condition of mental content, and consequently, a necessary condition of having a mind. Moreover, when the connection between perception and thought is studied, it becomes apparent that a link exists between humans as *psychological* beings and humans as *biological* beings. An account of the mind needs to recognise this link, especially when drawing conclusions to which this link is significant. For the personal identity theorist, this link is crucial to a correct understanding of personal identity. Minds are integrally linked to the world, and thus cannot be conceptualised in isolation from that world. Due to the link between personal identity and the mind, there is thus an inherent link between personal identity and the particular environments in which persons are situated and between their particular modes of interaction with that environment. It makes no sense to assume that personal identity can in any way be conceived of in isolation from the world in which persons live and interact. In addition to understanding the link between thought and the world, we also need to understand how thoughts are linked to each other. This topic is addressed in the next chapter.

Chapter 4 Psychological Coherence

4.1 Introduction

The importance to mental coherence of the relations between thoughts and the objects of thought has been emphasised. No less important to mental coherence are the relations between the thoughts themselves. Indeed, being related to other thoughts partly determines what a thought is. The psychological continuity criterion pays little attention to the relations between thoughts. For example, according to Parfit, personal identity is maintained in virtue of 'overlapping chains of strong connectedness,' regardless of how the mental items within these chains are connected. Without constraints on the causal connections involved, mental items could be connected by any means, such as by brain or memory transfers, reduplication, copying processes, and similar. Mental items that are malleable in this way must have a certain degree of autonomy and independence. More specifically, they cannot be substantially reliant on other mental items for their coherence and meaning. This position is sometimes referred to as *psychological atomism*. 'Psychological atomism' is the view that mental contents can stand alone, and do not depend for their nature, meaning, or significance on other mental items. Because they are independent in this way, complex thoughts, beliefs, and other mental items can be 'broken down' into smaller parts until the most basic psychological atoms or thoughts are reached (Honderich Ted 1995), p 64. According to Parfit, because they are self-sufficient and independent, these basic thoughts can be described without reference to the particular persons who own them, and so can be described *impersonally* (Parfit 1984), pp 210-225. In this chapter, I argue that psychological atomism and impersonal description are incorrect characterisations of the mind, as they represent neither the way that minds function and operate, nor the way that thoughts are experienced.

When psychological atomism and impersonal description are considered in more detail, it becomes evident that these two concepts are linked. To characterise mental items atomisticly is to hold that a particular mental item has meaning in its own right and does not rely for its coherence on other mental items. Because mental items do not rely for coherence on links with other mental items, it is incidental to which group or mental repertoire these mental items belong. This means that the particular ownership of mental items is not an important or decisive factor in their description or meaning. Whether mental items belong to one person or to another person makes no difference to the characterisation of the items concerned. It follows from this that the viewpoint from which an experience, memory, or belief is experienced is not a factor in the quality of that experience, memory or belief. Importantly, whether a particular experience is viewed from my viewpoint or from your viewpoint is irrelevant to the characterisation of the experience concerned. Such an experience can be described impersonally, that is, without specific reference to the particular person who has it.

Psychological atomism and impersonal description are correlative and mutually supportive. For example, given the requisite brain-transfer technology, one person could theoretically inherit one or two of another person's beliefs or memories, without having to inherit surrounding beliefs or memories. You could inherit the memory that you climbed Mount Everest, without inheriting the belief that there is such a thing as mountains. Your memory could be individuated and described impersonally, that is, without referring particularly to *you.* If your memory was owned by a different person, such as the person from whom you inherited it, it would be essentially qualitatively unaltered. Hence, according to this view, your memory is atomistic, and theoretically 'impersonal.'

A problem with this view is that it treats all beliefs and memories propositionally, that is, from a third-person perspective, while failing to allow for anything in addition to this. The belief that you climbed Mount Everest is similar in status to your belief that 2+2=4. There is no *personal* content in your belief that you climbed Mount Everest, no sense in which that belief is uniquely *your* belief, as opposed to someone else's belief. This view that all personally recalled beliefs and memories are no different to factual-type beliefs is problematic, as it does not accord with how many of these beliefs and memories are actually experienced. When I experience a memory of a particular incident, the memory is couched within the framework of my other memories, beliefs, and so on. It is the particular place of the memory within that framework which characterises the memory. There is a 'mineness' about it that cannot be reduplicated if the memory were 'transferred' to someone else. In fact, if the memory were part of a different framework, it would, by definition, be a different memory. But because psychological atomism does not recognise beliefs or memories as parts of frameworks, it cannot account for personal ownership.

Opposing psychological atomism is *psychological holism*. Psychological holism is the view that mental items gain meaning, at least in part, by virtue of their connections with other mental items. Thoughts and experiences are understood due to their connections with other thoughts and experiences. These, and other mental items, such as memories, form more or less coherent sets of mental items (or minds). Understanding some elements within these sets contributes to our ability to understand other elements in the same sets (Malpas 1999), p 79. Because psychological holism recognises the significance of a particular person's 'web of beliefs,' it supports the position that particular identities are relevant to the individuation and coherence of particular sets of beliefs and other mental states. Psychological holism and personal identification of mental states are thus correlative and mutually supportive.

Relation R involves commitment to psychological atomism and impersonal description. On this view, psychological connectedness is incidental to causal continuity, and the kind of network connections evident in psychological holism are not required to identify or describe mental items. As long as causal connections obtain between mental items, it is irrelevant whether these connections maintain coherence of content. Because items can be individuated in isolation from ownership, they can be 'described in an *impersonal* way' (Parfit 1984), p 210. This applies whether the items concerned are merely factual, or whether they are personally experienced memories. Psychological atomism and impersonal description are thus not only proposed by Relation R, but they are essential requirements of it. If psychological coherence were an essential feature of the psychological continuity criterion, it would be unable to hold to the Widest Criterion. This is because the Widest Criterion permits any cause of psychological continuity, regardless of whether that cause maintains psychological coherence. If psychological coherence was an explicit requirement of the psychological continuity criterion, it would be unable to commit to psychological continuity with any cause.

Consideration of the issues involved, however, indicates that psychological atomism is untenable. While rigid, tightly-knit connections are not always present in persons' belief systems or sets of memories, a certain degree of psychological coherence is necessary for mental states to be possible. A minimal degree of contentful connection between mental states is necessary for the states to have meaning. Unless there was some coherence between thoughts, no thoughts would be possible. How could one possibly have an isolated thought which was totally independent of any other thought? This is not plausible.[69] If there was no meaning between thoughts, minds would not be possible. These anomalies indicate that the psychological continuity criterion's commitment to psychological

[69]Even the thought 'I exist' cannot be isolated. What is 'I'? What does 'exist' mean? We must answer these questions to understand the thought, and must, therefore, have other thoughts also.

atomism and impersonal description is highly questionable. I argue that an atomistic, 'impersonal' characterisation of mental content misrepresents the mind's structure and mode of operation. I further argue that minds are predominantly holistic in structure and functioning and are individuated by ownership. To argue this case, I refer predominantly to memory, beginning by discussing *quasi-memory.* I then explicate further the positions of psychological atomism and psychological holism, and due to the importance of memory to this topic, follow this by examining several theoretical and empirical accounts of memory.

4.2 Quasi-Memory

A major impetus for the psychological continuity criterion's adherence to psychological atomism is the attempt to overcome the criterion's inherent circularity. First detected in Locke's theory by Butler,[70] the circularity in question here consists in the fact that to claim that a person's identity is grounded in that person's psychological states, presupposes that the states have already been identified as belonging to that person. Due to this presupposition, Butler concludes that memory could only provide *evidence* of personal identity, and could not be *constitutive* of it:

> But though consciousness of the past does thus ascertain our personal identity to ourselves, yet to say that it makes personal identity, or is necessary to our being the same persons, is to say, that a person has not existed a single moment, nor done one action, but what he can remember; indeed none but what he reflects upon. And one should really think it self-evident, that consciousness of personal identity presupposes, and therefore cannot constitute personal identity, any more than knowledge, in any other case, can constitute truth, which it presupposes (Butler 1867), p 194.

[70]The circularity objection was first mounted against Locke's account of personal identity in 1867 by Joseph Butler (Butler 1867), pp 193-198.

In contemporary terms, to say that what makes me the same person at t_2 that I was at t_1 is the fact that the psychological states I had at t_1 are continuously connected to the psychological states I have at t_2 is uninformative, since the claim that the states in question are mine presupposes that the states have already been identified belonging to me.

Parfit acknowledges that the psychological continuity criterion and the commitment to Relation R are vulnerable to the circularity problem. To overcome this problem, Parfit invokes the strategy of *quasi-memory*. Quasi-memory is unowned or impersonal memory. An account of a quasi-memory would include such details as times, events, places, and so on, but would not refer essentially to any particular individual person as the owner of that memory. Quasi-memory is a form of memory in which memories are like items of property, the ownership of which is independent of the property itself. For quasi-memory to work as a successful strategy requires that it comprehensively and accurately represents normal memory. If successful, its principles could be applied to all psychological states, thus furnishing a non-circular version of psychological continuity.

A detailed version of quasi-memory is found in Shoemaker's work (Shoemaker 1984). Shoemaker makes two claims concerning personal memory. First, he claims that for a person to remember an event requires that the person witnessed or experienced the event - this is designated as ‘the previous awareness condition.’ Second, he claims that a particular type of memory about oneself which is ‘full and accurate’ cannot be mistaken in respect of whether the person one remembers is oneself. These types of personal memories are not those made on the basis of observation, but rather on the basis of self-knowledge at the time the incident occurred. Thus, present-tense accuracy of knowledge is ‘preserved in memory’ (Shoemaker 1984), pp 20-21. Such memories are said to be ‘immune to error through misidentification.’

This second point can be understood more clearly by contrasting one's memories about oneself with one's memories about persons other than oneself. For example, although I seem to remember Frank telling me he had read a certain book, my memory could be mistaken. I may actually be remembering Bill, who is sometimes with Frank. The mistake could consist of either mis-identifying Frank in the first place, or by misremembering him. But, claims Shoemaker, in the case of some types of memories about ourselves, we cannot make this type of error. For example, if I correctly recall *myself* visiting my room at the university, checking my email, and returning home to begin re-reading Shoemaker's work on quasi-memory, I can be sure that I am not mistaking myself for someone else. In this type of memory, to remember is *just* to recall doing these things, and so to recall oneself doing them. Whereas to remember an incident concerning another person is to leave open the question of misremembering or mis-identifying, to remember correctly about oneself is to omit this possibility. Shoemaker claims this is not to say that one *could not* misremember incidents about oneself - clearly this is possible. Shoemaker gives the example of mistaking oneself in a mirror:

> If I say "I blushed when Jones made that remark" because I remember seeing in a mirror someone, whom I took (or now take) to be myself, blushing, it could turn out that my statement is false, not because my memory is in any way incomplete or inaccurate, but because the person I saw in a mirror was my identical twin or double (Shoemaker 1984), p 21.

The difference between the above case and one with immunity from error is that the former is a mediated experience (that is, I knew it was myself because I saw myself in a mirror), whereas the case of immunity from error is one in which the experience is direct and unmediated, (that is, I knew it was myself because I had self-awareness at the time, and was specifically recalling something about *myself*). In such a case, the immunity from error is maintained in memory:

> We might express this by saying that where the present-tense version of a judgement is immune to error through misidentification relative to the first-person pronouns contained in it, this immunity is *preserved* in memory (Shoemaker 1984), p 21.

That is to say, I might mis-identify John as the person I saw yesterday, but not that is was me who remembers seeing him (Shoemaker 1984). Shoemaker then claims that interpreting one's apparent memories as knowledge of one's own past experiences may seem trivially analytic, but need not be so. If one person could seem to have direct knowledge of experiences had by another person, then we could consider these 'memories' in isolation from the experiences that gave rise to them. Thus, if we could possess the knowledge of another person's experiences, such that the apparent recall appears to one in the same way as the memories of one's *own* experiences, this would constitute *quasi-memory knowledge.* Having quasi-memory knowledge would mean that a cognitive state qualitatively identical to the cognitive state generated by the original event existed in someone, but not necessarily in the same person who had the original experience. A person with such knowledge could be said to *quasi-remember:*

> Let us speak of such knowledge, supposing for the moment that it is possible, as "quasi-memory knowledge," and let us say that a person who has this sort of knowledge of a past event "quasi-remembers" that past event (Shoemaker 1984), p 24.

In the case of a quasi-memory then, a previous awareness condition would be present in someone, but not necessarily in the person who remembers the event. An example of this might be that I have a quasi-memory of attending a lecture by Frank on Plato. This would mean that I have the present cognitive state with this content, but that it may not have been me with the cognitive state at the time of the lecture. It may have been a person other than myself who attended, and who had that earlier state. Without the initial cognitive state, I do not satisfy the previous awareness condition, and, in this case, although my memories would *seem* to be of me, there is no certainty that they would be:

> Quasi-memory, unlike memory, does not preserve immunity to error through misidentification relative to the first-person pronouns (Shoemaker 1984), p 28.

Shoemaker claims that in this universe we do not have quasi-memories, but that the notion of quasi-memory prompts the identification of certain features of memories which permit the circularity objection of the psychological continuity criterion to be surmounted. When considering the possibility of quasi-memory, Shoemaker adds the requirement that memories are correctly causally connected to the events they represent, which for him means that a present cognitive state occurs due to its causal link (by an *M-type* causal chain) to a previous cognitive state.[71] Thus, while in this universe, ordinary memories are correctly annexed to the events they represent due to the owner's previous awareness condition, in another universe, a correct causal connection could be of some other kind, such as by means of an M-type causal chain (Shoemaker 1984).

Shoemaker's earlier thought experiment provides an example of an M-type causal chain.[72] If the brain of Brown were to be transplanted into the skull of Robinson, and 'Robinson' were to subsequently display the memories, character dispositions, and so forth, which were formerly attributed to Brown, then, according to the psychological continuity criterion, the former person 'Robinson' would have become the former person Brown, to then be referred to as ‘Brownson.’ But if we assume this identification merely because Brownson has Brown’s memories, Shoemaker claims we would be engaging in the same sort of circular argument which troubled Butler. That is, we know that Brownson is Brown because he has the same memories which Brown formerly had, and what identified those memories as Brown’s is simply the fact that they were his. To overcome this circularity, we could look to the causal connection, that is, an M-

---

[71] Shoemaker claims that an M-type causal chain is one that entails a quasi-memory which is causally linked to the corresponding past cognitive and sensory state of which it is a memory. Causal links do not have to be of the 'normal' sort, and could involve partial or complete brain or memory transfers (Shoemaker 1984), p 35.

[72] As referred to in Chapter 2.

type causal chain. We could then claim that Brownson is the former Brown, not just because he has qualitatively similar memories to the former Brown, but because these memories are *correctly causally connected* to Brownson by means of the brain transfer operation. According to this hypothesis, the causal connection is correct because the previous awareness condition exists in *someone,* although it is not necessary that this someone is the same person as the one who apparently recalls the original experiences. Thus, in this case, the causal connection is not dependent on the identity of the rememberer, and, therefore, Shoemaker claims, the circularity problem is overcome (Shoemaker 1984).

To appreciate how quasi-memory features in the debate about personal identity, we need to consider its impact on first-person memories by imagining that we inhabit a quasi-memory possible universe. Here, I could not take for granted that any first-person memories which I appear to have are mine merely because I appear to have them. I may appear to remember myself sitting at the computer writing this thesis, but whereas in the present universe, this first-person knowledge would, due to the previous awareness condition, be immune to error through misidentification, in a universe with quasi-memory this would not be so. In such a universe, instead of the previous awareness condition which guarantees my memories in this universe, I would be relying on the presence of a non-branching M-type causal chain, which, of course, could never be confirmed. So I could never know if 'my' memories were mine. Also, one could never know 'who' one was! In other words, Cartesian scepticism is now a possibility, even for self-identity.

In the case of the imagined Brownson story, Shoemaker claims that although Brownson does not inherit his traits in the normal way, the fact that Brown formerly had these traits was an important part of the cause of Brownson now having them, and that further:

> It is only where we suppose that the traits of things at different times are causally related in this way that we are entitled to take the similarity of something at one time and something at another time as evidence of identity (Shoemaker 1984), pp 44-45.

This means that if the identity of the rememberer is not part of the description of the memory, identity could then be ascribed on the basis of the appropriate causal links. In other words, part of what it is for me to know that I am the same person at t_2 that I was at t_1 is for me to have a memory which is correctly causally linked, that is by a non-branching causal chain, to a former event. By imagining a universe in which memories can be transferred from one person to another, Shoemaker claims that although we are unable to perform the required transplant operations, the fact that we can *imagine* them happening entitles us to employ the projected consequences in personal identity arguments. By means of quasi-memory, we can consider a person's memories in isolation from that person. Rather than accepting that my apparent memories are really mine because they have the requisite previous awareness conditions, they are mine because they have the correct causal links. Shoemaker refers to these different interpretations respectively as 'strong' (that is, memory which entails a previous awareness condition in the rememberer), and 'weak' (that is, memory based on an M-type causal chain) remembering, designated as 'remember$_{S}$' and 'remember$_{W}$'. He claims that in actual remembering, it is difficult to know which of these senses of remember reflects our common understanding. Further, it is only because we do not really have branching causal chains that we have no need to make this distinction. It would seem that for the memory criteria of personal identity to work, an understanding of memory in the 'weak' sense is essential, although Shoemaker is a little ambiguous on this point:

> But I do not think that this question is especially important. We can defend the spirit of the claim that memory is a criterion of personal identity without settling this question, although in order to defend the letter of the claim we must maintain that in its ordinary use "remember" means "remember$_{W}$' (Shoemaker 1984), p 43.

It seems, however, that how we understand remembering *is* important, and that the latter point is crucial. For if the above claim is wrong, and my memories are mine in virtue of being causally connected specifically to me by means of the previous awareness condition, rather than because they have some other non-standard type of causal connection, such as an M-type causal chain, then the original circularity objection remains unanswered.[73] We need to bear in mind that the difference between Shoemaker's two kinds of remembering is the difference between memories identifiable by means of a previous awareness condition in the rememberer, and memories identifiable by means of an M-type causal chain. Memories based on a previous awareness condition in the rememberer are memories which are standardly causally connected to the person who had the original experience, which means the experiencer and the rememberer are one and the same person. Memories based on an M-type causal chain are not necessarily causally connected to the person who had the original experience, they could be causally connected to a different person, in which case, experiencer and rememberer are not the same person (Shoemaker 1984), p 36. This means that for quasi-memory to work as part of a non-circular account of personal identity, an adequate characterisation of the pertinent memories cannot include additional information, such as reference to other psychological states, or to the body of the rememberer, as such information would inevitably entail reference to already established ownership. If the ownership of the memory is part of what we understand a memory to *be,* then its place in a psychological continuity criterion of personal identity cannot be impersonal, and consequently, the circularity present in this criterion cannot be overcome by recourse to quasi-memory (Shoemaker 1984) . The very notion of quasi-memory would involve a vindication of what we do indeed take memory to be.

[73]While the 'previous awareness condition' has some merit (although, as is shown in Section 4.6, not total merit), it turns out that reference to this condition is not the prime method of justifying the presence

4.3 Psychological Atomism

Although Parfit presents a simplified version of quasi-memory, its underlying principles are essentially the same as those of Shoemaker's version:

> (1) I seem to remember having an experience,
> (2) *someone* did have this experience,
> and
> (3) my apparent memory is causally dependent, in the right kind of way, on that past experience (Parfit 1984), p 220.

In other words, my memories are not necessarily mine due to the normal causal connections between *my* experience and *my* memory, but are mine due to the 'right kind of causal connection,' which could be any kind whatever. Because quasi-memory permits any type of causal connection between memories, it involves psychological atomism. That is, memories can be understood on an individual basis, rather than as parts of semantic or epistemic networks.

Parfit claims that hypothetical examples provide a basis for the evaluation of quasi-memory, as the consideration of 'certain imaginary cases' assists the understanding of 'actual people in ordinary lives' (Parfit 1984), p 219. To evaluate quasi-memory requires ascertaining whether these 'imaginary cases' are analogous to actual life. As an example of quasi-memory, Parfit presents a thought-experiment in which part of a brain is transferred from one person to another person. In the scenario envisaged in this experiment, brain biology has advanced to the stage where individual memories can be located in particular brain parts, while neurosurgery has developed to the point where these minute parts can be physically transferred from one brain to another. Such transfers allow one person to quasi-remember another person's past experiences. In Parfit's story, part of Paul's brain is transferred to Jane. Jane subsequently 'remembers' events which occurred during Paul's trip to Venice (Parfit 1984), p 220. Other than

---

of authentic memory in personal identity accounts. As will be shown shortly, the holistic structure of memory is more important.

the brain connections involved, these memories have no particular connections with Jane's other memories.

Jane quasi-remembers the events 'contained' in the transferred memories. These events include walking on marble paving, hearing the flapping of pigeons' wings and the cries of gulls, seeing lightening, and shaving. As quasi-memories, Jane experiences these memories in the first-person mode, although their content and quality are the same as they would be if these memories were experienced by Paul. Jane is unaware that 'her' memories are 'Paul's' memories unless he informs her about the memory transfer, or unless the memories are obviously anomalous, such as the memory of shaving. Apart from these two provisos, Jane's experience of Paul's memories is indistinguishable from her experience of her own memories (Parfit 1984), p 221.

From this scenario, Parfit draws conclusions about the operation of genuine memory experiences. He claims that the scenario shows that causal connections between memories do not need to rely on ownership, but rather on overlapping strands of connected experiences. Although connections within and between strands are required, the *kind* of connection is irrelevant. These overlapping strands can thus provide the necessary connectedness in a person's life, without presupposing the person's identity:

> Overlapping strands of strong connectedness provide *continuity of quasi-memory.* Revising Locke, we claim that the unity of each person's life is in part created by this continuity. We are not now appealing to a concept that presupposes personal identity. Since the continuity of quasi-memory does not presuppose personal identity, it may be part of what constitutes personal identity. It may be part of what makes me now and myself at other times one and the same person (I say 'part' because our criterion also appeals to the other kinds of psychological continuity) (Parfit 1984), p 222.

The unity of a person's life is explained in terms of 'overlapping chains of strong connectedness' of impersonal memories. Because these memories are impersonal, they do not presuppose personal identity, and therefore can be used as part of the criteria for personal identity. Consequently, claims Parfit, the account of personal identity is no longer circular.

The above scenario, however, is rife with problems concerning the meaning and interpretation of memory. A mental item is treated, on this account, as a memory, regardless of the type of connection it holds with the original event. Memory is traditionally held to be a form of recall.[74] This recall is normally understood to be from the perspective of the original experiencer. It is this perspective which characterises the experience, and consequently, the memory of the experience. Yet, according to Parfit, although Jane has not experienced the original events, she experiences the memories of these events in qualitatively the same way as a person who *has* experienced such events. But, without the experiential connection to the original events, it is not clear to me how Jane's mental events can be memories. As Jane did not actually experience the events which the memories are about, there is little, if anything, to distinguish her mental events from delusions. A delusion is: 'a false belief.' When Jane 'remembers' events experienced by Paul, she is either experiencing a mental state which she knows is not her own, or she is having the false belief that the mental state *is* her own, in

[74]Some different types of definitions bring out the intent of memory to refer to some type of recall. A typical *commonplace* definition characterises memory as: 'the faculty by which things are recalled to or kept in mind; to recall is: to summon back; bring back to memory, recollect, remember' (Turner 1978), p 429. A typical *philosophical* definition characterises memory as: '1. The mental function of retaining information about stimuli, events, images, ideas, etc. after the original stimuli are no longer present. 2.The hypothesised "storage system" in the mind/brain that holds this information. 3.The information so retained' (Reber 1985), p 429.

And finally, a typical *psychological* definition characterises memory as: '1.The recall of past experiences that are assumed to be stored in unidentified biological structures, most probably neural. 2.Remembering, that is, behavioral sequences in which ongoing, prior, and later experiences are synthesized. 3.The treatment of past experiences in a manner analogous to electronic information processing' (Popplestone and McPherson 1988), p 234. While these definitions *could* be understood to refer only to a brain state at the time of recall, regardless of the connection between the memory and the original event, such an understanding goes against the traditional meaning of memory, and thus should be explained and justified.

which case she is having a delusion. In neither case does her experience accord with what is normally understood as memory.[75]

Parfit's scenario also misconstrues the very nature of the memory process. For him, memories are just 'experienced' without any specified background. Jane's memories of incidents experienced by Paul are supposedly qualitatively the same memories which Paul would experience. This view treats memory as a discrete item, which can be passed around from one person to another, without any substantial change being made. It assumes that single, unattached memories could be experienced and understood. Such memories do not have to belong to epistemic or semantic networks to have meaning. In Jane's case, she can remember items regardless of their epistemic links, or lack of such links. Her 'memories' of Venice do not depend on whether she knows what or where Venice is. We are not told how she recognises Venice, or what knowledge she must already posses in order for her memory to make sense. She must 'remember' the items and experiences contained in the transferred memory, regardless of whether she is familiar with them. But without the appropriate supporting framework of experiences and ideas, the memories could be totally mysterious. It is difficult to conceive of a memory which could stand in isolation from other memories or mental items, such as beliefs or plans. Unless it was connected to such items, a memory is unlikely to make any sense. In taking an atomistic view of memory, Parfit and similar theorists violate the very nature of the memory process.

According to Simon Beck, thought experiments that discuss the transfer of brains or brain parts simply assume that psychological states have the necessary

[75]Consider the situation in which you vividly describe an event in your life to me. I may imagine it, rehearse it, and accurately recall the details, maybe complete with mental imagery. But I was not present at the event, and so cannot be considered to be remembering it. I am merely remembering what you have told me. This must not be confused with reconstructed memories which will be dealt with shortly.

atomistic character for 'transfers' to be possible. But according to Beck, this view of mental states is mistaken. Beck claims it is undeniable that mental states do not actually operate in this discrete fashion, but can only be what they are do to their connections with other mental states:

> The holistic character of our mental states like beliefs and desires seems undeniable. Particular beliefs and desires function only against the background of a vast network of other beliefs, desires, and so on. And thus we have no reason whatsoever to believe that a single mental state, or even a set of mental states, could be removed from an individual's psychology, leaving the remainder intact (Beck 1991), p 4.

Beck claims that this conception of mental states denies not only mental holism, but also physicalist theories, such as type-identity theories. These theories are not built on an indiscriminate characterisation of identity, but on the identity which holds between particular individuals at particular times. Thus, if particular thoughts are taken out of the context in which they are situated, there is no reason to assume that these thoughts would, or even could be the same if they were 'transplanted' elsewhere. Beck argues that there is no way we could never know how particular mental states and particular portions of physical matter relate to each other. He claims that it is unwarranted to assume that transfer or reduplication of physical matter inevitably entails the replication of particular mental states, as there 'is no reason to believe that the transplanting of a small amount of brain matter from one individual to another will bring with it a particular mental state' (Beck 1991), pp 1-4. Beck argues that it is simply not enough to say that we are, as yet, unable to perform such atomistic type operations on humans, as the holistic character of mental states prevails, regardless of the physical matter out of which such states are 'constructed'.

Another problem with quasi-memory is that it undermines and misconstrues the relation between memory ownership and personal identity. If we accept that quasi-memories are legitimate memories, the psychological continuity criterion

itself would be rendered vacuous. Were technology to advance as Parfit envisages, and memory transfers between persons become a reality, then psychological connectedness could refer to any artificially constructed and potentially modifiable series of supposed memories, rather than to the memories of incidents experienced by a specific, individual person. This means that the life or supposed memories of a single person might never be accurately represented in a single memory chain. It could never be assumed that an individual's 'memories' referred to his own life. The memories could equally apply to incidents experienced by several different individuals.

As a means of explicating personal identity, the psychological continuity criterion would thus be uninformative and completely impotent. There are, moreover, a number of points on which Parfit's account is simply confused. On the one hand, the transferred memories are referred to as belonging to someone - in one sense they are Paul's, though they are also Jane's - and yet on the other hand, they are also held to be unowned or impersonal. Marya Schechtman argues that memories cannot be treated in this arbitrary way:

> simply deleting the “nametag” from a memory is not sufficient to make it nondelusional, and that in order to make an apparent memory truly nondelusional one will either have to presuppose the identity of the rememberer with the person who had the experience, or else remove so much of the content of the memory that it is no longer plausible to say that what is relevant to personal identity in genuine memory is preserved in quasi memory (Schechtman 1990), p 79.

Schechtman contrasts Parfit's memory description with that of Edward Casey's account of a family visit to the cinema.[76] When details of the outing are recalled, they occur with a myriad of other little details and anecdotes, such as the demeanour of the children, their responses, and also other memories. Some memories are clear and definite, while others are indistinct and uncertain. The

[76]Schechtman takes her account of memory from (Casey 1987), pp 25-26.

memories are not presented as discrete items of experience, but rather as elements in an intricate network, with complex and dynamic connections to other mental states. They have a richness and depth which is absent from quasi-memory (Schechtman 1990) (Schechtman 1990), p 81-82.

Schechtman imagines what might happen if Jane were given the memories in Casey's account. Without their accompanying background details, these memories might be unrecognisable, disconnected, alien, or incoherent, just 'a blur of unidentifiable sights and sounds' (Schechtman 1990), pp 82-83. But if the memories did include their connections with other mental states, confusion would arise. These connections would fail to synchronise with those of Jane's own mental states, making integration between these alien memories and Jane's own life history problematic. The transfer of the whole of a person's memories would not solve the issue, since this would give rise to a question as to whether the transfer in question was a transfer merely of memory or of a whole person. The identity of the memories themselves would then be problematic, as it would be unclear which had priority, the identity of the memories or the identity of the person who has them. And, unless the memories were tagged to a particular person, as Schechtman notes, the distinction between the memories and delusions would be unclear:

> What this discussion has shown is that, on either of the two possible pictures of what it is to have a quasi-state, quasi-states fail to do the work they are supposed to because they include either too little or too much of the state they reproduce. If they include too little, they do not capture what is relevant to personal identity, and if they include too much, then, unless sameness of person is assumed, they are delusional (Schechtman 1990), p 86.

If too much is transferred with the memory, a whole person seems to be transferred, but the means of identifying that person is inherently problematic. On the other hand, if too little is transferred, the transfer itself is meaningless and incoherent.

In summary, quasi-memory is a thoroughly unsatisfactory strategy. The discontinuity between quasi memory and other mental states implies that quasi memories would be unintelligible. Quasi-memory is so unlike normal memory that it offers no help in unravelling the issues concerning memory and personal identity.

4.4 Psychological Holism

In contrast to the atomistic view of quasi-memory, psychological holism recognises that mental states are necessarily united to other mental states. It recognises that while mental states can, in a sense, be considered individually, their unity with other mental states is a necessary condition of their coherence. For many theorists, this raises the question of what underpins mental unity itself. David Hume was particularly perplexed with this question, as he could detect no underlying self in which to ground such unity. He concluded that mental states were united into bundles, but could detect no principle of unity, other than that of the association of the ideas within particular bundles.[77] Hume later expressed dissatisfaction with this solution, but was unable to offer a better one (Hume 1962). Many contemporary theorists have questioned the idea that resemblance is, in itself, sufficient to explain the connections between mental states.[78]

---

[77]Hume claimed that, while some philosophers claimed they could perceive a self, he could not. He consequently made his claim: 'I may venture to affirm to the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement' (Hume 1888), 1.4.6.

[78]For example, Barry Stroud asks what if there are a group of perceptions which *do* resemble each other, occurring of, for example, the Eiffel Tower - are we to assume that they are all necessarily present in a single mind - could they not be occurring in several different minds? (Stroud 1977), pp 124-125. Conversely, there could be a multitude of *different* perceptions simultaneously present in a single mind? In this case, there may be no obvious connection between the ideas generated. (I might simultaneously be thinking about my thesis, and about what I will cook for dinner tonight. There is no similarity in the ideas involved). David Pears also criticises Hume's view of psychological unity. He claims that Hume's account gives rise to the problem of loose 'psychological integration ... like a group of buildings around a farmyard' (Pears 1990), p 122. Without a persistent self, the mind of a person 'is a sequence of ever-changing impressions and ideas' (Pears 1990), p 124.

Pears argues that unless there are principles which bind them together psychologically, there seems to be no reason why particular ideas are unified in minds the way we normally experience them, or expect them to be. Another problem pointed to by Pears is that Hume's account does not adequately explain what distinguishes one particular bundle from another. Pears uses the example of an outbuilding, which may be identifiable without knowing to what complex or other building it belongs. There is no evident principle or rule which determines which buildings belong to one group rather than to another. Pears maintains that Hume presents mental items as being like this, but that in reality they are not. It makes no sense, according to Pears, to talk of them without attributing ownership (Pears 1990), p 122.

The idea of mental unity is also a central issue in the Kantian attempt to explore the conditions for the possibility of experience. In Kant, this idea of unity is primarily expressed through the idea of the unity of apperception. This unity can be understood in terms of the ability for the self-ascription of mental states that is, the ability, as Kant suggests, for mental states to be understood as one's own:

> It must be possible for the 'I think' to accompany all my representations; for otherwise something would be represented in me which could not be thought at all, and that is equivalent to saying that the representation would be impossible, or at least would be nothing to me (Kant 1929), B 132.

The possibility of *synthesis,* or the mind's combining activity, is directly tied to such apperceptive unity. Synthesis is the combining of representations, as delivered by the faculty of sensibility, in relation to concepts, derived from the understanding, so as to constitute a single experience or judgement. Such combination is achieved through the referring of representations back to a single 'I.' We can say, then, that mental unity is required, since the having of mental states itself depends on a certain connectedness obtaining between those states, and between the elements of those states. The unity of states and the unity of the 'I' involved in self-ascription, that is, the unity of apperception, are thus one and the same (Kant 1929), A 84-130; B 117-169.

There remains an ambiguity, however, as to how we should understand the nature of such apperceptive unity, and the nature of the synthesising activity with which it is associated. Addressing this issue, Malpas claims that synthesis is underpinned by embodied agency, and that it is this that ultimately unifies mental content. Malpas argues that only a predominantly holistic account of mental states can explain how mental content and action are possible.

According to Malpas, experiencing subjects, minds, and actions are integrally related and mutually constitutive. While subjects exist, they are not constituted

prior to experience, but rather in tandem with experience. An essential part of any experience and mental content is that a certain form of self-awareness is part of the experience itself. For example, it is a formal requirement of me seeing the hedge outside my window that I am aware that it is *me* seeing the hedge, rather than you seeing it. My awareness does not occur by my first having the experience, and then appropriating it to myself, as if I could equally have appropriated it to someone else. I do not first see the hedge and then decide it is me seeing it. My seeing the hedge is, by definition, my experience. Thus, having the experience and knowing it is me who has it are simultaneous. This point is not undermined by mistaken experiences. Even my mistaken experiences are, by definition, my experiences.[79] Awareness that an experience is mine is integral to my awareness of myself. Put another way, one does not grasp oneself as a subject separately from grasping oneself as an experiencing subject. Knowing oneself as a subject is integral to knowing oneself as an experiencing subject (Malpas 1999), pp 72-78.

Malpas also notes that for a person to have experience requires that coherent connections obtain between the experiences of the person concerned. Rather than being regarded atomisticly, the experiences of a single person must be considered as part of a coherent framework of experiences. Only through their connections with a variety of mental items, can any mental items have sense or meaning. For example, knowing it is a green hedge outside my window entails knowing about the existence of plants and colours. To believe that one was once in Russia requires believing that the world is divided into countries, one of which is currently called Russia. Even mistaken beliefs about being in Russia cannot be held in isolation from the requisite background of associated beliefs. Recognising the connections between our experiences provides the sense of ourselves as subjects persisting through time. It should be noted here that the requirement of

[79]For example, I may mistake the hedge for a real one when it is artificial, or I may mistake the yellow leaves for flowers.

meaningful connections between mental states does not mean that every mental state is connected to every other mental state by means of its content, although clearly certain minimal such connections are necessary for states to have meaning (Malpas 1999), pp 78-80.

A key feature of mental holism, for Malpas, is the relation between mental content and action. Mental states are organised in relation to action, while the ordering of action is seen only in relation to mental states. Hence, an integral relation exists between the two. This can be understood by considering the difference between action, which is (in general) rationally motivated,[80] and behaviour, which could be an uninitiated bodily response (such as shivering). Actions which are rationally motivated draw on beliefs and other mental states. For example, because I believe that water quenches thirst, I drink water when I am thirsty. Without the belief, the action would not occur. And without the framework of beliefs within which the belief about water sits, I could not have such a belief.[81] Beliefs are parts of schemas (or small systems) contained in larger networks, forming a person's total belief system. The particular concatenation of elements within a given circumstance impels the relevant belief or desire to motivate action. These elements are the state and circumstances of the world in which an agent is located, combined with the agent's particular aggregated network of mental content. According to circumstances, certain elements of the network are drawn towards the requisite action:

> Like a spider's web in which a fly has been caught, and in which all the threads have pressure exerted on them from a single point, so mental states are similarly organised in relation to, or are 'pulled towards' the current activity of the agent (Malpas 1999), p 96.

---

[80]Of course, not *all* action is rationally motivated, but those which are not are contingent on a world in which most action is, in some way, rationally intended. Unless this were the case, the distinction between rational and non-rational actions could not be made.

[81]This should not be confused with an instinct to drink water, such as might be had by an animal. I am referring here to full cognitive awareness. Of course, creatures can be motivated to drink water without this awareness, but this is not what is being discussed here.

Action brings together the relevant elements of mental content, providing a focus for their unity. While integration occurs at the point of action, this does not necessarily entail all parts of a person's mental framework or belief system. This is not required for action to occur, as all parts of the framework are not relevant to all actions (Malpas 1999), pp 96-99.

For some theorists, such as Jerry Fodor, however, involvement with one element in a holistic system necessarily means involvement with them all. He claims that this involvement could potentially destabilise the whole system, as it means that if one element in the system were changed, then *all* elements would change. The ability of speakers to share meaning would be compromised. An utterance derives its meaning from the constituent parts of a person's belief system. Fodor argues that understanding one component of the system is dependent on understanding it all. Therefore, unless a speaker understands the total system of another, she cannot understand any of it. Without this understanding, argues Fodor, the communications and meanings of one person may not be comprehended by another person. Similarly, if holism operates within different systems of thought, it could seem that no two disparate systems of thought, such as two scientific theories, could refer to the same objects, such as stars. Because little of the beliefs of these disparate systems (such as ancient Greek astronomy and modern astronomy) are held in common, then neither are their objects of reference. The problem then arises as to whether beliefs between different persons and systems can be identical, merely similar, or indeed, if they can have anything in common at all (Fodor and Lepore 1992), pp 4-22.

Other theorists, however, such as Bilgrami and Malpas, believe that these objections can be met, without holism being compromised. Bilgrami, for instance, claims that holism must be constrained by externalism. This means that not all beliefs are called into play at the point of action, but only those that are relevant to

the particular action being performed. According to Bilgrami, beliefs are affected by external causes only to the degree to which these beliefs are implicated at any one time. Not all beliefs are always implicated, due to the presence of two levels at which they can be conceived. The first is the conceptual level, referred to as a the 'meaning-theoretic level.' This level is an aggregation of a particular person's mental repertoire. The second level is that of content, referred to as 'local,' and concerns beliefs selected from the theoretical level, which are relevant to any particular experiential situation. The theoretical level is general, while the local is specific, the latter being that at which action occurs. Because beliefs extracted at the local level are only those implicated in any given situation, a complete change in holistic content does not occur.

Bilgrami gives the example of two persons drinking water to quench their thirst. Both have the belief that the water will perform this task. One of the persons, skilled in chemistry, also holds beliefs about the water's chemical composition. However, the belief that the chemical composition of water is H_2O contributes nothing to the action of drinking the water. Hence, when the action is performed, the beliefs about the water's thirst quenching ability are invoked at the local level, while those concerning its chemical composition are not. Thus, while the aggregated level of beliefs is different for each person drinking water, those at the local level are the same. Were the beliefs about the water's chemical composition to change, it would not automatically follow that the beliefs about its water-quenching ability would also change. Because the degree to which beliefs are brought into play at the point of action is limited, the outcome feared by Fodor is curbed. It therefore does not follow that all beliefs are subject to inevitable change. Thus, Fodor's claim that persons could neither maintain beliefs, nor share the same beliefs is unsustainable. Similarly, the fear that beliefs and theories about beliefs is threatened, is insupportable (Bilgrami 1992), pp 142-150.

In summary, the holistic account of mental states implies that important semantic and epistemological relations pertain between different mental states, but that these relations are not inherently destabilised every time a belief or mental state is changed. The key point is that contentful relations between mental states are essential for those states to have any meaning or content at all.

To see how well these views relate to the way mental states are actually experienced, some theoretical accounts of memory are now considered. These accounts demonstrate that memory is, in principle, a holistic process, in which memories do not occur in isolation from other mental states, but indeed, are tied to, and contributed to by those other mental states. Empirical accounts of memory also support this claim.

4.5 Memory Theories

'Memory' is broadly understood as the recall of earlier events, or as a form of stored knowledge.[82] This knowledge is derived from events experienced by a particular person, and therefore holds the perspective of that particular person. Were the memories of one person to be experienced by a different person, the original perspective of the memories could not be duplicated, as it is not possible for one person to have the same perspective on the world as another. Some further distinctions concerning memory are important.[83] A genuine memory is the (more or less) accurate recall of an event which actually occurred. Genuine event memories are causally connected to, and in some ways like, the original event. *Mistaken* memories are memories of events that did occur, but are different in some important or significant way to the original event. *False* memories are mental events that are taken to be memories, but which are not causally connected to any event which actually occurred. False memories include errors,

---

[82]For example: 'The ability of the mind to store and recall past sensations, thoughts, knowledge etc' (Wilkes and Krebs 1995), p 830.

[83]See also (Schechtman 1990), p 78.

hallucinations, delusions, and one form of imagination.[84] Consideration of genuine and faulty memories brings out the significances of these distinctions.[85]

Plato and Aristotle were among the first thinkers to seriously consider memory. They saw memory as a 'trace' in the mind, although its exact nature could not be determined. Augustine took memory to be a reproduction of an original event, which had been 'stored away' in the mind. Don Locke initially held memory to be the retention of ideas in the mind, but later considered this problematic as the perceptions on which the ideas were based no longer existed.[86] He subsequently saw memory as the revival of earlier perceptions (Locke D. 1971)[87], pp 3-6. The *Representative* and *Realist* theories of memory arose from these early views. The Representative theory of memory holds that memory is a mental state which represents a past experience.[88] The representation is in the form of an image in

---

[84]It is important to be aware of possible ambiguity when referring to imagination, especially when associating the term with memory. An imagined event can be of two kinds: first, an event which I mistakenly claim to remember, but which has not occurred, I only imagined that it did; in this case I am not aware (until and unless it is drawn to my attention) that I am imagining it. For example, I may strongly intend to return a book to the library, and then imagine that I have done so. It this case I make a mistake (I make the distinction between imagination and error here as they are not necessarily the same kind of mistake. An error could be of the form of mis-remembering *that* you have done something, such as returning the book to the library. Imagination could be the misremembering/ imagining of the actual incident itself). The second type of event is one in which I deliberately and knowingly imagine something, such as a plot for a story, or how I will spend the next holidays.

[85]Memory, its operation and its aberrations is a large and complex topic, and can clearly not be completely covered here. I have therefore selected material for these sections to give a general overview of the field, and to bring focus to areas which impact on psychological continuity and quasi-memory.

[86]Based on his view that the mind had no innate ideas, Locke saw memory as 'the keeping of those simple ideas which from sensation or reflection it has received' (Locke 1959), 2.10.2. Memory is thus an essentially mental operation. Memory is needed because persons are unable to foreground all their ideas at the same time, and thus require somewhere to keep them: a kind of 'storehouse' where ideas reside until invoked at some later time. Subsequently, Locke modified his view to the claim that because ideas are actually 'perceptions in the mind', they could not exist beyond the duration of the original perception. A memory, therefore, is the activity of reviving an earlier perception. Locke's 'storehouse' view was modified in a later edition of the *Essay.*

The paradox of possible 'unperceived perceptions' prompted his addition: 'But, our *ideas,* being nothing but actual perceptions in the mind, will cease to be anything when there is no perception of them; this laying up of our ideas in the repository of the memory signifies no more but this,- that the mind has a power in many cases to revive perceptions which it has once had, with this additional perception annexed to them, that *it has had them before.* And in this sense it is that our ideas are said to be in our memories, when indeed they are actually nowhere;- but only there is an ability in the mind when it will to revive them again, and as it were paint them anew on itself, though some with more, some with less difficulty; some more lively, and others more obscurely' (Locke 1959), 1.2.10.2.

[87] To avoid confusion, John Locke is referred to throughout at 'Locke' and the more recent Don Locke is referred to as 'Locke D.'

[88] The following sections are based on work from (Locke 1971).

the mind. In some sense, the image reproduces the past experience, causing memory knowledge. However, to view memory as merely an image leaves open the question as to the distinction between genuine memory images, mistaken memory images, and images which are completely false. Attempted solutions to this problem by Hume and Russell actually help little. Hume claims that memory can be detected because it is more vivid than imagination. This is not necessarily the case, as imagination frequently outstrips memory in vivacity and intensity. Russell distinguishes memory from imagination by claiming that genuine memory experiences are accompanied by feelings of 'familiarity' and 'pastness.' As with Hume's distinction, however, these feelings could equally apply to an imagined event, and tend to presuppose the memory experience they are trying to justify (Locke D. 1971), pp 7-16. Don Locke claims that the fundamental distinction between memory and imagination is that memory refers to an event *which actually occurred,* whereas imagination does not. This means that a genuine memory is one which is causally connected in the right kind of way to an event which took place at a time prior to the occurrence of the memory.[89] The problem is less one of defining what memory is, but rather more of knowing how to make the distinction between genuine and non-genuine memories:

> The Representative theorist finds himself imprisoned within his imagery, with no way of confirming that the imagery does reveal the past, as it has to if we are to have memory-knowledge (Locke D 1971), p 17.

According to Thomas Reid, a more direct theory was needed, and thus the *Realist* theory of memory arose (Locke D. 1971), p 20.

The Realist theory of memory holds that a memory experience is a 'direct awareness of the past' as opposed to a mere representation of it (Locke D. 1971),

[89]By this I mean that the causal connection between the prior event and the memory is one of recall, as opposed to some other kind of causal connection, such as a hypothetical form of 'memory transplant.' This means that if I have a thought of an event which actually occurred, but which was not causally connected to the original event by means of recall, my thought would be one of (coincidental) imagination rather than memory.

p 21. Memories are not new items, but rather the re-appearance of old ones. For the representativist, the mental occurrence called a 'memory' is numerically distinct from the original mental occurrence; for the realist, it is numerically the same (Locke D. 1971), pp 24-26. Nevertheless, the Realist theory incurs similar problems to the Representative theory. Distinctions are difficult to make between genuine and non-genuine memories or between deteriorated or distorted memories. Verification of memories is still dependent on factors internal to the memory experience, rather than on factors outside it (Locke D. 1971), pp 29-31. These difficulties mirror those discussed in the previous chapter, and show that a different approach to memory is needed.

Don Locke notes that both the above theories hold in common the following three points:

1 Remembering is an occurrence, something we *do;*
2 The occurrence consists of an experience, such as an image;
3 Based on the occurrence, we know various facts

(Locke D. 1971), p 32.

These three points, however, do not apply to all memories. Not all memories entail remembering particular incidents or having mental images. Don Locke cites remembering facts such as the election results. This seems to be a completely different act to remembering, say, where the car is parked. And neither necessarily entails remembering any particular act. Remembering where I parked the car may involve remembering the act of parking it, but it may not. I may simply remember where the car is. These realisations led to the reappraisal of memory and to the recognition that it can refer to remembering incidents, or it can refer to knowing facts. Don Locke distinguishes *Factual Memory, Practical Memory,* and *Personal Memory* (Locke D. 1971), pp 32-77.

Factual memory refers to knowing specific facts, or 'retained factual knowledge' such as scientific facts, the days of the week, or the date of one's birthday (Locke

D. 1971), pp 52-52. Practical Memory refers to abilities and skills, such as being able to drive a car or play a musical instrument (Locke D. 1971), pp 64-65. Personal Memory refers to memories of one's own experiences (Locke D. 1971), pp 70-71. In the course of research, these basic ideas have developed and expanded into more refined taxonomies. As the research below demonstrates, these taxonomies show that connections between memories and other psychological states are a crucial factor in the ability to understand and experience memories, and they therefore support the claim that memory is, in principle, a holistic mental process. This argues against the atomistic characterisation of mental states, and more specifically, against the possibility of quasi-memory.

4.6 Memory Research

Beginning in the late 1900s, much memory research focused on understanding the memory process, and identifying its main influential factors.[90] Early work, such as that of Ebbinghaus, examined the duration of memories.[91] Recording his own attempts to remember 'nonsense syllables,' he concluded that much learned material is soon forgotten. More recent studies indicate that a steep 'forgetting curve' results when learned items are meaningless (Weiten 1992), p 251. Results are more positive when items with meanings and contexts are committed to memory. This indicates that meaning and context are relevant to memory. Items isolated from contexts are very difficult to remember, thus posing a challenge to the atomistic view of memory.

As research developed, diverse traditions were set up by Galton and Freud. Galton's interest was in the kinds of associations which assist memory, such as verbal or sensory cues (Robinson 1986), pp 19-23. Freud's focus was on the role

---

[90] This section is drawn from a variety of sources, including (Goethals and Solomon 1989), pp 2-11, and also from the works listed in footnotes below.

[91] Ebbinghaus' seminal article on memory research *On Memory* was published in 1885.

of memory in neurosis, and thus on subliminal causal links in recall (Rubin 1986), p 20. Binet's research on children's memories[92] reveals that considerable reconstructive process occurs in memory (Goethals and Solomon 1989), pp 2-3. Reconstruction is where memories become conflated with other memories and experiences. Exact sequences of events may not be remembered, but may be combined with other recollections. Studies conducted by Bartlett indicate the importance to reconstruction of a person's prior knowledge and expectations.[93] These trends in research reveal the significance to memory of other mental items, thus supporting the trend towards a more holistic approach.

Neurological progress prompted some researchers, such as Lashley to try to locate memory traces in the brain.[94] His failure to do so led to the conclusion that memory was distributed throughout the brain, rather than localised in a specific brain part (Goethals and Solomon 1989), p 3. This view was modified in 1984, when McCormack and Thomson's experiments on rabbits revealed a small area of conditioned response. They studied the behaviour of rabbits, before and after inserting a small lesion into part of the brain. This enabled them to locate the site of the conditioned response, proving that particular brain parts were involved in particular memories. Although sites like this are sometimes located, many memories are still held to be distributed throughout the brain. The following taxonomies are the outcome of study in recent years.

4.6.1 Memory Taxonomies

Two basic kinds of memory are *procedural memory* and *declarative memory.* Procedural memory refers to learned abilities and skills, which we may or may not remember learning, but which we can execute without undue concentration, such

---

[92]Binet's study: *Binet & Henri* was published in 1894.

[93]Bartlett's ideas were presented in *Remembering: A Study in Experimental and Social Psychology* was published in 1932.

[94]These traces were referred to 'engrams', and were a major focus for Karl Lashley in the 1950s.

as typing. Closely connected to procedural memory is *implicit memory.* This is sub-conscious memory, which can only be accessed indirectly, and may be reached when treating amnesia or other disorders (Weiten 1992), pp 258-259. Both procedural memory and implicit memory are related to contexts, and are thus difficult to conceive of in isolation from these contexts. Because contexts are composed of experiences and other mental states, the atomistic conception of memory fails to adequately characterise these types of memories.

Declarative, or explicit memory refers to factual information remembered explicitly. It includes *semantic memory* and *episodic memory.* Semantic memory refers to general facts such as mathematical knowledge, or the letters of the alphabet. Episodic, event, or personal memory is memory of particular events one has experienced (Weiten 1992), p 259.

Personal memory can be further divided to distinguish between memories which pertain specifically to life events which concern oneself, and those which do not. For example, 'rote learning' is a form of personal memory. So is witnessing an event in someone else's life, such as a wedding or a birthday, or even just seeing an anonymous car drive down the road. But unless these events involve one in some meaningful way, they are not significant to one's own life events. Memories which do involve oneself in a personal and significant way are those that compose a person's own life-story, or 'autobiography.' These memories are referred to as 'autobiographical memories.' Autobiographical memories have special relevance for personal identity, as they contribute to a person's self-conception. The sense of a continuing self over time is provided by the connecting of a person's life-events within a continuing history (Brewer 1986), p 33. It is often the more significant or public of these events by which a person is identified by others, such as into which family one is born, what job one has, one's address, partner, children, achievements, and so on. Because many of these things are inherently

connected, it is difficult to conceive how memories relating to them could be understood in isolation from other memories or mental states, and thus, how an atomistic account of memory could be sustained.

Personal or episodic memory operates in three basic stages, encoding, storage and recall. Encoding is the forming of a memory code as the event is 'put' into memory. Some kind of connection is made with other items in the memory store. Codes can be acoustic or semantic, and affect the effectiveness of recall. Poor encoding means poor recall, while failure to encode means failure to recall (Weiten 1992), pp 238-246.

Storage can be sensory, short-term or long-term. Sensory storage is the initial reception of sensory information by one of more of the sensory modalities. It is controlled by attention and lasts three to four seconds. Unless encoded, information is lost. If encoded, it passes to short-term storage for about ten to twenty seconds. Storage is prolonged by rehearsing and combining with other mental contents, during which it is in 'working memory.' Long-term storage is the retention of information over extended periods, lasting from days to years, or even lifetimes. Information may be retained in semantic form, specific propositions, or images. Three models are used to explain storage:

- The *connectionist* model, in which memories are connected by networks of association clustered in the brain.
- The *spreading activation* model, in which information entering one node spreads to other nodes, connecting up associated thoughts, ideas, and memories. This model may also be understood as a *semantic network*, which could be complete, or composed of smaller networks, comprising grouped concepts in the form of *scripts* or *schemas.*
- The *hierarchical organisation model,* in which memories are stored in relation to their place in a graded structure, such as by being part of a concept, which in turn is part of a larger concept, and so on.

It is likely that all models apply to some degree, depending on the type of memory and circumstance (Weiten 1992), pp 238-246. These models support the view that

memories are not isolated items of experience, but are defined in part by their connections with other memories and associated mental states.

Recall of memory (retrieval) is complex, and subject to several factors. Key to effective retrieval is adequate encoding, and circumstances conducive to recall. Retrieval does not normally occur in abstraction from surrounding events, experiences and concerns. These items provide cues which activate recall. Failure of recall could be due to either poor encoding, or to inadequate and inappropriate recall cues. This means that some memories could be permanent, but virtually inaccessible. Retrieval can be enhanced by activities, such as citing places and events, posing suggestions under hypnosis, re-enacting situations and moods and similar activities. These devices contribute to recall, often invoking a reconstructive process, in which missing factors, forgotten from the original experience, are supplied. The longer the lapse between original event and recall, the more reconstruction occurs (Weiten 1992), pp 247-251.

4.6.2 Reconstruction

Reconstruction is subject to a variety of influences. Some examples bring out ways in which these influences affect what is recalled and how it is recalled. A case of recall involving considerable reconstruction was that studied by Ulric Neisser, of the subject John Dean.[95] Testimony was given by Dean concerning meetings held over several years, both with Nixon, and with members of Nixon's administration.[96] Unbeknown to Dean, the conversations held at these meetings were recorded. It was thus possible to compare Dean's testimony with these recordings. Dean admits that he used newspaper clippings of the events in question to aid his recall. When his testimony was later compared with the tapes of the events, it was discovered that some facts were conflated, and many details

[95]For the full details of this study, see (Neisser 1982).

[96]As counsel to the President of the United States of America in the early 1970s, Dean testified in 1973 before the 'Watergate' Committee of the US Senate concerning apparent covert government involvement in burglary and deception (Neisser 1982), p 140.

were mixed-up. Conversations which took place on one occasion were combined with parts which occurred on others. But what was apparent was that the general *gist* of the testimony was correct. Neisser claims that Dean's testimony, while not verbatim, was:

> … not altogether wrong. On the contrary, there is a sense in which he was altogether right; a level at which he was telling the truth about the Nixon White House. And sometimes, ... he was more specifically right as well (Neisser 1982), pp 157-158.

While the memory of specific events was unclear, the underlying semantics of those events was retained. That is, although the details may have been inaccurate, the basic meaning behind them was not. As a result of the close correlation between the *gist* of Dean's testimony, and the *facts* of the tapes, Nixon had to resign, and many high-ranking staff of the White House were sent to gaol. Neisser concludes that many memories operate in this reconstructed fashion. Minute details are lost or misremembered, and sequences of events merged with other similar sequences. The overall coherence of events is, however, maintained. These types of memories are active and dynamic, as opposed to being static, photographic-like representations of original events (Neisser 1982), pp 139-159.

4.6.3 Eye-witness Testimony

Studies of eye-witness testimony also demonstrate the presence of reconstruction during recall. A review of such studies, conducted both inside and outside the laboratory, indicates that 'pure recall' of witnessed events is unlikely. A variety of factors influence recall, often undetected by witnesses. For example, how much a witness can retain is influenced by exposure time at the original event. Laboratory tests which showed images for varying amounts of time revealed a correlation between the degree of retention and the amount of exposure. Also influential is the amount of time that elapses between exposure and recall. Memories can be affected by day-to-day activities, which can cause deterioration or forgetfulness.

Other factors, especially in non-laboratory situations, are distance, poor lighting, crowds, colour-blindness, and other distractions such as personal stress. In both real-life crime situations and laboratory tests, stress is a considerably influential factor, due to its effect on elements such as adrenalin levels, heart rates, and other bodily functions. These elements influence concentration and focus. To show how this might happen, researchers give the example of subjects who were denied food for over 24 hours, and their subsequent tendency to 'see' food in blurred images on a screen (Buckout 1982), pp 116-119.

Social factors are also shown to influence recall. This was detected in studies of selections made from line-ups in photographs. Tests in which subtle suggestions were put from superiors to subordinates showed a high percentage of response to these suggestions in the choices made. In other studies, the choice of particular words, such as 'smashed' or 'hit' demonstrated the influence these words had on the request to estimate speeds.[97] Similarly, a test in which six out of seven people (falsely) claimed that the longer of two lines was really the shorter, influenced the seventh person to agree with them.[98] Other studies indicated that many witnesses become more certain about their claims as they progressed from their first police interview, through the ranks to the grand-jury, indicating that the witness had 'filled in' the former uncertainties (Buckout 1982), pp 120-122. These studies revealed, not only the unreliability of eyewitness testimony, but also that plausible reasons were often available to explain discrepancies. The studies also reveal that factors adjacent to memory experience are influential in the construction and experience of the memory concerned.

---

[97]A study by Elizabeth Loftus showed that the word 'smashed' drew higher speed estimates that the word 'hit' when questioning viewers about an automobile accident film (Buckout 1982), p 120.

[98]This test was designed to test the relevance of conformity among persons. In the test, six out of the seven participants pre-arranged to give the wrong estimate of line length. In spite of the obvious mistake, the 'naive' person more often than not agreed with them (Buckout 1982), p 122.

4.6.4 Flashbulb Memories

A more specific form of recall is that referred to as 'flashbulb' memories. These memories are so-called as they are supposedly illuminated by some intense event or experience to which they are attached. Research indicates that flashbulb memories can be generated by incidents of private concern, such as personal shocks, or public concern, such as assassinations or attempted assignations. A remarkable example of the phenomenon concerns the moment when persons heard of the assassination of President John F. Kennedy in 1963.[99] According to a study by Brown and Kulik, this incident generated more flashbulb memories than any other incident studied at the same time.[100] Their study showed that many people 'captured' certain details of what they were doing at the precise moment they heard the news. These details are preserved with the memory of Kennedy's assassination, and are consequently subject to the same vivid recall (Brown and Kulik 1982), pp 23-25.

Of interest to the researchers was why the Kennedy incident was so generative of flashbulb memories. Brown and Kulik compiled relevant data. Their study covered several assassinations, and drew on separate groups of black and white Americans. By comparing the relative significance of the victims to the well-being of the groups, a correlation was found between the needs, concerns, and survival of the groups and the amount of flashbulb memories held in relation to particular victims. For example, a much higher proportion of black than white Americans held flashbulb memories for the death of Malcolm X and Martin Luther King.[101] The number of flashbulb memories held for the death of Kennedy was almost equal - 39 out of 40 whites, and 40 out of 40 blacks. Whereas, the deaths of the

---

[99]For the full details of this study, see (Brown and Kulik 1982).

[100]This needs to be kept in context. We can imagine similar phenomena applying at other times and places, such as when ancient emperors, or persons like Lenin died, or maybe in more recent times, Princess Diana.

[101]Both were associated with working for the civil rights of black Americans.

black leaders was considerably more significant to the black population than to the white, the death of Kennedy was perceived as being equally significant to both.

Also relevant was the type of detail recalled. It usually included place and activity, but not minor or insignificant details. In most instances, the memory had been rehearsed, either privately, or by retelling, indicating that an amount of reconstruction was involved. Brown and Kulik considered these facts in the light of the envisaged evolutionary significance of memory. This significance concerns the likely connection between memory selectivity in primitive humans, and the imperative to remember the relevant facts about predators and other dangers, such as location, type of threat, and so on. A similar significance is mirrored in the above study. The stronger the connection between the victim and the rememberer's well-being, the stronger and more prevalent the memory of factors related to the incident. Brown and Kulik conclude that there is a connection between the flashbulb phenomenon and survival or well-being of the rememberer (Brown and Kulik 1982), pp 28-40. These points make sense, as they represent an extreme example of an aspect which often accompanies most memories, namely a correlation between a memory's importance and the quality of its recall.[102] Reflection on the above study indicates the possibility that this correlation has evolutionary origins.

4.6.5 Summary Of Memory Research

The above studies and considerations of the memory process indicate that memories are not inert and discrete, but rather are dynamic and contextual. They do not occur 'out of nowhere,' but belong to a framework within which they make sense.[103] The various components of the framework, such as background

---

[102]The correlation is clearly influenced by whether the knowledge was present in the initial situation. The quality of recall by a witness who *knew* of the importance of an event at the time the event occurred may be considerably better than that of a witness who only learned of the importance later.

[103]Of course we do sometimes have 'flashes' of memory, which come unbidden and may surprise us by their unexpectedness. But our lives would make no sense to us if all memories were like this. I am here discussing the normal occurrence of memory, as it is this which forms the background of

conditions, biological, social and personal concerns influence the degree and type of reconstruction that occurs. This supports the view that memories could not be arbitrarily removed from their contexts and placed elsewhere, and then be assumed to remain intact, as if they were still in their 'home' locations. The holistic structure of memory does not fulfil the requirements of quasi-memory, but rather indicates that memories can only by understood in relation to the context in which they are situated. This conclusion is also supported by studies of false memories.

4.7 False Memories

False memories are memories which are totally fictitious, and refer to supposed events which have never occurred. Although reconstruction occurs in the course of normal memory, reconstruction is not synonymous with the experience of false memories. Reconstruction and degeneration are a normal part of the memory spectrum. This does not mean, however, that *anything* can be construed as a memory. There is a distinction between a memory which is more or less genuine, and an apparent memory which is completely false. This distinction is important to understand, as it is relevant to Parfit's argument regarding transferred memories and ownership of memories. If we were to agree that 'transferred' memories or supposedly 'impersonal' memories were acceptable as genuine memories, we would be failing to distinguish between genuine and false memories, such as has happened in the case of Jane's memories of Venice. The distinction between genuine and false memories is a real one and needs to be recognised. If the distinction between genuine and false memories were overlooked, one would never know which memories applied to oneself, and thus what one had done, or, more importantly, who one really was. As examples of false memories, 'memory illusions' and 'false recognition' are considered. It will be shown from these

---

coherence which is necessary for such unexpected memories to be possible. Unless we had such coherence, we would not recognise the 'out of the blue' occurrences.

considerations that these types of false memories do not arise out of nowhere, but, like genuine memories, have contentful connections with other mental states.

4.7.1 Memory Illusions

The phenomenon of 'memory illusions' was investigated by Lampinen et al.[104] Memory illusions are one form of 'false beliefs.' A memory illusion is defined as an occurrence in which false memories are experienced phenomenologically, or as: 'an incorrect belief about the past that is experienced as a memory' (Lampinen, Neuschatz et al. 1998), p 183. The subject falsely remembers an event with the same perceptual awareness as that of genuine memory experiences.[105] By contrast, a false belief is the incorrect belief *that* something happened, regardless of whether the event itself is recalled. The researchers' concern was to explore the nature of memory illusions, as these challenge our notion of first-person verification of memory experiences. When I remember myself doing something, I want to claim that the fact that I remember the incident vividly warrants my memory of the incident being taken seriously. In most instances, my claim would be uncontroversial, but there are cases where it would not be. Mistakes can be made in remembering, posing the question of what, in addition to personal recall, is required to verify a memory experience as being genuine? Lampinen et al. see this as a crucially important question, due to its relevance to cases of reported child abuse in the case of purportedly recovered memories. To address this issue, Lampinen et al. devised investigatory techniques, which recognise that subjective experience is relevant to memory illusions, but that subjective experience is, in principle, unobservable, and therefore, difficult to study (Lampinen, Neuschatz et al. 1998), pp 181-183.

---

[104]For full details of the investigation, see: (Lampinen, Neuschatz et al. 1998). For their comprehensive list of references, see particularly pp 219-224. As the study is extensive, this section is unable to cover all the researchers' points, draws on those aspects most relevant to the present work.

[105]An example of this would be Jane 'remembering' Paul's experiences.

Lampinen et al. investigated memory illusions by examining many studies of false memory reports. By examining a range of studies, they produced some illuminating general principles about the operation of memory. A key initial task was to differentiate between memory illusions and propositional-type false beliefs. This is important, as it locates the areas in which the study of phenomenological characteristics is required. Studies were usually conducted by providing test situations in which subjects were presented with events, and then asked to recall them under a variety of manipulated conditions.

One study concerned incorrect 'recognition' of words in word-lists. Continuous lists of words were given to subjects, with the instruction to nominate words already presented. Errors were made when words were similar, or semantically connected. For example, if the word 'sleep' was presented, some subjects falsely reported that it had been presented earlier, when in fact the earlier word was 'bed.' Similar studies in which word 'lures' were presented to subjects, yielded incorrect reporting.[106] When 'nurse, sick, lawyer, medicine' and so on were presented, subjects falsely reported the word 'doctor.' Another study, in which non-famous and famous names were mixed in particular ways, yielded mis-reporting of non-famous names as being famous. Some researchers recognised and extended Bartlett's work on memory reconstruction, by conducting similar studies to those by Bartlett.[107] In one example, subjects were given varying amounts of information concerning certain characters, such as Helen Keller or Adolf Hitler. In instances where subjects *knew* the identity of the characters, more mistakes were made concerning other information given in the study, than was made by subjects to whom the identity was not given. The falsely 'recognised' information was information which would cohere with what was generally known about the character, but which was not actually supplied in the study. Another

---

[106]'Lures' are specifically referred to by these and other researchers, rather than 'cues' which may have been expected. This is probably because the words are deliberate ploys, designed to mislead.

[107]Bartlett's studies were mentioned earlier in this chapter, in Section 4.6.

range of studies concerned simple and complex sentences. By supplying subjects with a range of sentence components, false combinations were reported as having been given. Lampinen et al. concluded that false recognition ranges from simple to complex situations: 'subjects will not only falsely recognize and recall simple stimuli such as words and names but also more complex stimuli such as stories and sentences' (Lampinen, Neuschatz et al. 1998), pp 183-186.

These studies led to the production of a specific three-staged formula, found to produce a high percentage of false reports.[108] Stage one is where an event or film of an event is witnessed. Stage two is where misleading or neutral information is given to some subjects. Stage three is where all subjects are questioned regarding the original event. In one example, subjects were first shown a slide presentation of a vehicle accident. Some subjects were then given a questionnaire containing misleading information, in which a yield sign was substituted for a stop sign. When all subjects were subsequently questioned, it was found that 75% of non-misled subjects gave correct answers regarding the signs, whereas only 41% of misled subjects gave correct answers. The study demonstrates how post event information can mislead and distort memory (Lampinen, Neuschatz et al. 1998), p 187.

Even warnings about inaccurate information does not necessarily prevent such information from contaminating a memory. In one investigation, although subjects were told that their information had been manipulated, test results were not significantly affected. The warnings did not prevent the subjects concerned from being influenced by the false information. For example, in one eyewitness post-event study, subjects were warned that police information might be inaccurate. Contrary to expectations, this did not produce a higher accuracy-rate of reporting in warned subjects than in unwarned subjects. Similar results have been found in

[108]This methodology was devised by Elizabeth Loftus in the 1970s, and, due to its success, has been used by many researchers since then (Lampinen, Neuschatz et al. 1998), p 187.

many different types of tests, such as word-list tests which had both pre and post-event warnings. Researchers concluded that warnings specifically prevent neither false recognition nor false recall. They infer from this that memories resilient to warnings are likely to be illusions rather than inferences, as inferences would be more vulnerable to change when warnings were given (Lampinen, Neuschatz et al. 1998), pp 202-204. The significance of this type of study is that in spite of warnings, memories can be unavoidably contaminated by their surrounding circumstances.

Another study showed how repeated suggestions can produce a completely false 'memory.' The researchers, along with the subject's brother, conducted conversations over several days, with the result that the subject eventually 'remembered' an event in which he had been lost in a shopping mall as a young child. The memory was described as 'vivid.' No such incident had ever occurred. Another study showed how related items can be falsely reported. In a case where many items of a particular kind are placed in a room, reports of the items often include related items which are not present. For example, a room containing office items was reported to contain books when none were present (Lampinen, Neuschatz et al. 1998), pp 187-188. These, and many similar studies, show that false memories can and do occur. To discover why, several models of memory were examined.

The *Perception/Reception* model recognises the similarity between perception and memory, and draws conclusions about memory by comparing the two. Tests which concerned reports of perception and memory found that in neither case did subjects treat them as 'discrete' experiences. Responses to events were coloured by the subjects' own general knowledge and personal concerns, raising doubt in the appropriateness of the 'storehouse' model of memory (Lampinen, Neuschatz et al. 1998), p 189.

The *Fuzzy Trace* theory holds that memories are encoded in both *verbatim* and *gist* forms. 'Verbatim' refers to specific information and to its surface format. 'Gist' refers to the relevant information in a more generalised form. The theory holds that genuine memories can be produced by either, whereas false memories are produced by the latter only. In cases of verbatim recall, experience is richer than gist recall. But where verbatim traces are not rehearsed, they can decompose, and become integrated with other gist representations, producing traces which are a mixture of both, thereby 'creating' false memories (Lampinen, Neuschatz et al. 1998), pp 190-191.

The *Source Monitoring Framework* theory holds that memory illusions result from a mistaken conception of the origins of memories. The mistake is to assume that the memories have been generated from 'without,' rather than from 'within.' Certain types of inferences which lack emotional or contextual content can 'overlap' with inferences which contain more perceptual and emotional content, thus causing confusion to the rememberer regarding the origin or 'source' of the memory (Lampinen, Neuschatz et al. 1998), p 191. By investigating whether the source of a particular memory is a 'recalled' experience, or a held belief, source monitoring tests can ascertain whether a false memory is phenomenologically experienced, or merely held propositionally. The limitations of subjective reports are recognised, resulting in some researchers augmenting these tests with others which yield more detail (Lampinen, Neuschatz et al. 1998), pp 200-201.

Lampinen et al. concluded that the above results imply that the memory process is an interrelated process, rather than a discrete one. Memory consists of interconnected components within a framework, each of which, to some degree, depends on other components.[109] Lampinen et al. argue that while individual

[109]For instance, Lampinen et al. cite short-term memory, attentional capacity, output interference, long-term memory, memory strength, and so on. While unobservable, they are known to exist due to behaviours which *are* observable, and due to the explanatory power which these components provide (Lampinen, Neuschatz et al. 1998), p 193.

memory components cannot be studied directly, they *can* be studied by observation of the framework within which they sit:

> It is perfectly reasonable to draw conclusions based on theoretical frameworks to the extent that the framework as a whole is subject to confirmation or disconfirmation (Lampinen, Neuschatz et al. 1998), p 193.

Memory as a *framework,* rather than as a unitary operation, permits the study of memory illusions, and of the conditions conducive to their occurrence. This supports the view that an atomistic conception of memory is not only false, but, given the results of the various studies above, also contrary to the basic principles of the memory process itself.

For Lampinen et al., the value of the above investigation lies in its ability to inform and predict the future direction required for further memory illusion study. But for the present thesis, the interest lies in the investigation's support for the contention that memories, in principle, operate in a predominantly holistic fashion. The above studies reveal ways in which memories can be influenced and coerced by adjacent mental experiences. The studies show that lures can influence persons to mistakenly believe they have already seen previously unpresented words. They also show that the manipulated presentation of false material can cause the creation of completely false memories. These false memories are not discrete, but are conditioned by personal circumstances and concerns. Gist representations can be inaccurately remembered, thereby confusing the rememberer regarding the source of memories. Gist can also mistakenly be reconstructed and experienced phenomenologically as an episodic memory, deceiving the rememberer regarding the validity of the memory. Memory illusions appear to be an extreme and distorted version of the reconstructive process which occurs in normal memory. Illusions appear to be produced by the faulty amalgamation of various mental processes, rather than by a single, discrete act. These problems show why the verification of a particular memory often involves reference to the

mental framework to which the memory belongs, rather than just to the individual memory itself. The way that false memories occur further supports the holistic conception of memory, and adds more weight to the case against the possibility of quasi-memory. The view that memory is an integrated and interrelated process is further supported by investigations of false recognition, across different groups of subjects.

4.7.2 False Recognition

Amnesia researchers investigate the causes of different types of amnesia. Factors which assist this research concern rates of false and true recognition. Due to damage in medial regions of the temporal lobes and associated structures in the diencephalon, amnesiacs have impaired short-term memory.[110] To determine precisely the effect of this damage to memory, recognition tests have been carried out. These tests involved comparing recognition responses between amnesiacs and matched control groups. Early tests differed in methodology from later tests. Results between earlier and later tests concerning the rate of false recognition of the different groups differed also. A third test was carried out, to try to isolate the variable factors which led to these contrasting results (Schacter, Verfaellie et al. 1998), pp 668-669.

A general aim of all tests was to detect the rates of true recognition, absence of recognition, and false recognition.[111] 'False recognition' is the mistaken claim that an item has been previously presented, when the former item was merely similar. A difference between earlier and later tests was the rate at which this occurred. In the first test, amnesiacs displayed a higher rate of false recognition than did controls. In the later tests, the rate of false recognition of amnesiacs was *less* than

---

[110]The temporal lobes are two of the four lobe divisions of the brain, and are located laterally on each side; the diencephalon is the part of the brain located between the two cerebral hemispheres and the midbrain (Tortora 1980), pp 407; 615.

[111]As it is false recognition which most concerns the present inquiry, reference to the other results will not be made. The study is extremely complex, and too lengthy to deal with in full here. Referring to one section of the results only in no way distorts that section, or the results as a whole.

that of controls. All tests comprised the presentation of words lists, with the instruction to differentiate between words which were presented for the first time, and words that were presented previously. Subjects comprised controls and amnesiacs. The tests included semantically related lures, which, in this case, were words not identical with previously presented words, but were words semantically related to them (Schacter, Verfaellie et al. 1998), 669-670.

The first test concerned controls and Korsakoff amnesiacs.[112] It comprised one continuous word list, in which some words were presented once only, and other words presented more than once. The subjects' task was to read down the list, indicating whether a word was being repeated, or presented for the first time. New words were 'signalled' by the presentation of either an associated word, a homophone, or a synonym.[113] If subjects were distracted by these lures, they would falsely 'recognise' the new word as having already been presented. On the first presentation of a word, subjects designated 'new,' and on subsequent presentations, designated 'old.' The subjects' scores thus represented the degree of correct and false recognition that occurred. The results of the studies showed that the amnesiacs had a higher level of false recognition than the controls (Schacter, Verfaellie et al. 1998), p 669.

The second test concerned controls, Korsakoff amnesiacs and non-Korsakoff amnesiacs. It consisted of two word lists. The first was studied with the aim of memorising as much of it as possible. The second list comprised the recognition test. Subjects had to indicate which words on the second list had already appeared on the first list. As before, lures were presented. In this test, however,

---

[112]Korsakoff syndrome is a psychological disorder caused by alcoholism, and involves 'mental confusion, hallucinations, and memory losses' (Weiten 1992), p 182. It was thought that Korsakoff amnesiacs might differ from non-Korsakoff amnesiacs by having damage to frontal lobe areas. This extra damage may increase the type of memory deficit experienced, and was thus of interest in the third test.

[113]Associated words might be 'food, eat'; a 'homophone' is a similarly sounding word, such as 'brake, break'; a 'synonym' is a similarly meaning word, such as 'find, locate'.

amnesiacs showed *less* false recognition than did controls. Schacter et al. sought to discover the factors behind the variation in results between the earlier and the later tests (Schacter, Verfaellie et al. 1998), p 669.

The researchers conjectured that the variation might be attributable to the different range of subjects, or to the different details of tests, or possibly both. Subjects in the first test comprised controls and Korsakoff amnesiacs only. Subjects in the subsequent tests also included non-Korsakoff amnesiacs. This was relevant, as the extra damage in the frontal lobe suffered by Korsakoff amnesiacs may have been responsible for the poorer retention rate of already presented words. If this were the case, it would explain why the results of the earlier and later tests were different. The different details between earlier and later tests concerned the way in which words were presented for recognition. In the first study, new words were signalled by individual lures. Subjects had to remember the individual lures to distinguish between the lures and the newly-presented words. For controls, this involved remembering single items, which could feasibly be done by means of episodic memory. This would be manageable for controls, but due to their incomplete memory mechanism, difficult for amnesiacs. Hence, controls showed a better rate (that is, less) of false recognition than amnesiacs. In the second test, however, the words to be repeated were presented on a preliminary list, rather than 'mixed in' with the test words. Hence, the ability to detect a repeated word depended, not on a lure which had just been presented separately, but on remembering whether a lure had been presented on a previous list. For controls, the previous list would have been too much to encode as individual episodic memories, and would therefore have been predominantly encoded as 'gist.' But gist, being less specific than episodic memory, would be more vulnerable to deception from lures, and thus more likely to generate false recognition, such as for example, mismatching a later word 'food' with the earlier word 'eat.' Due to impaired functioning, amnesiacs encode less gist than healthy

controls, and therefore have less gist available to mismatch. The task then, for Schacter et al., was to find out the cause of the discrepancy between the results of the earlier and later tests (Schacter, Verfaellie et al. 1998), pp 669-670.

To investigate this issue, Schacter et al. devised a third test. Locating the relevant differences between the two earlier studies required recognising the above variation in both subjects and test details. Strategies were devised to illuminate the differences in performance between non-Korsakoff and Korsakoff amnesiacs, and the differences caused by test variations. Subjects were Korsakoff amnesiacs, non-Korsakoff amnesiacs, and matched controls. The strategy involved the initial presentation of six study lists. Following the study of these lists, subjects were presented with a recognition test, which included new words, words from the study list, semantically-related lures, and non-semantically related lures. The procedure of studying lists, followed by testing was repeated, with a total of five tests being conducted (Schacter, Verfaellie et al. 1998), pp 669-670.

Schacter et al. predicted that repeated exposure to the lists would facilitate the operation of episodic memory in controls, producing decreasing rates of false recognition in each test. They predicted that the defective memory mechanisms of Korsakoff and non-Korsakoff amnesiacs would inhibit the use of episodic memory, thereby preventing the progressive improvement of false recognition rates. They reasoned that false recognition in amnesiacs resulted from the degeneration of semantic gist, rather than from unsuccessful attempts to use episodic memory. If this was correct, tests which permitted the build-up of episodic memory in healthy controls, but which, due to the impairment of memory mechanisms, could not cause the same build-up in amnesiacs, would show that the rates of false recognition would decrease (that is, improve) in controls, but would fail to decrease in amnesiacs. Schacter et al. conjectured that the extra damage to frontal lobe areas in Korsakoff amnesiacs, as opposed to non-Korsakoff

amnesiacs, might be also relevant. If these differences were relevant, they would show up by Korsakoff amnesiacs having an initially high rate of false recognition in comparison to controls, rather than merely one which deteriorated across trials (Schacter, Verfaellie et al. 1998), pp 669-670.

The results from the tests verified much of what had been proposed. Non-Korsakoff amnesiacs initially showed less false recognition than Korsakoff amnesiacs, although this difference lessened as the trials progressed. The levels of false recognition of non-Korsakoff amnesiacs steadily increased across trials, resulting in a crossover with the rate of Korsakoff amnesiacs. The increasing levels of false recognition in Korsakoff amnesiacs was attributed to the deterioration of cumulative gist representations. Mixed results comprising non-Korsakoff and Korsakoff amnesiacs showed them to initially have slightly less false recognition than healthy controls. Mixed groups of non-Korsakoff and Korsakoff amnesiacs showed less improvement in false recognition across trials than healthy controls. The difference in results between non-Korsakoff and Korsakoff amnesiacs was probably caused by several factors and requires more testing, although the extra damage to frontal lobe areas in Korsakoff amnesiacs was probably a factor. Healthy controls showed decreasing levels of false recognition across trials. This was attributed to cumulative improvements in explicit recollection as trials progressed. In general, data was consistent with that of previous trials conducted by Schacter et al. concerning rates of reduced false recognition in both amnesic subgroups.

As in the previous studies, the impact of these results for researchers lies in the results' predictive power for future research, but for the present discussion, the results are relevant to the coherence of memory. They show that the coherence of particular memories is tied to, and constrained by factors which, while adjacent to particular memory experiences, are yet integral to them. Results concerning all

three groups brought out the difference in the episodic memory operation of each. In each case, factors outside the particular memory were found to be influential. For example, healthy control groups demonstrated less false recognition across trials due to the increased rehearsal of mental items. The improved episodic memory counteracted the influence of lures, resulting in improved recognition rates. Non-Korsakoff amnesiacs demonstrated increased rates of false recognition due to the retention of gist, which, influenced by impaired recall, misled them - much like the case of naive healthy controls during the initial tests. And finally, the initially higher level of false recognition in Korsakoff than in non-Korsakoff amnesiacs demonstrated their impaired memory functions, regardless of contamination, either by lures or faultily recalled gist. In each instance, memory operation was influenced and circumscribed by the conditions in which the memory experience occurred. In no instance did the memories appear to operate discretely, nor were they unaffected by other memories, or other circumstantial influences.

These test results show that even misremembering is not an atomistic, unattached mental occurrence, but is one which is influenced by surrounding and influential factors. The results show that even false memories are attached to other memories and other mental states. It appears that not only are genuine memories unlike quasi-memories, but that false memories are unlike them also.

Like genuine memories, false memories demonstrate the holistic nature of memory, and its ability to be affected, influenced, and in some instances, obstructed by the circumstances and concerns within which it occurs. Indeed, the above studies indicate that not only is memory influenced by the surrounding circumstances and concerns of the rememberer, but that these circumstances and concerns are often *part* of the memory itself. These findings provide further evidence against the atomistic conception of memory entailed in quasi-memory and the possibility that memories could be characterised in the way required by

the psychological continuity criterion, particularly as it relates to Relation R. If the commitment to Relation R is dependent on the success of quasi-memory, it is now evident that this commitment cannot be sustained.

4.8 Summary

This chapter has assessed Relation R's commitment to an atomistic and impersonal characterisation of mental states. This commitment is necessary to meet Relation R's causal requirements, namely, psychological continuity with any cause. Psychological continuity with any cause is only possible on an atomistic characterisation of mental states, as normal causation requires reference to other mental items, and ultimately to ownership. I have argued that the characterisation of mental states required by Relation R misrepresents the mind's structure and mode of operation and consequently is unsustainable. I have also argued that minds are predominantly holistic in structure and functioning, and correspondingly, are individuated by ownership. To defend my claims, I have explored psychological atomism and psychological holism from both theoretical and empirical perspectives. This has involved a consideration of quasi-memory and an investigation of genuine memory.

The exploration of quasi-memory shows it to be a faulty strategy, as it fails to represent the structure and operation of genuine memory experiences. Quasi-memory fails to cohere, either with other memories, or with the experiences from which these 'memories' first arise. Work considered from Davidson and Malpas shows how the intelligibility of a particular mental state or memory depends, at least in part, on its place in an integrated network of other similar mental states. It is only by belonging to that network that the state has meaning and can be identified as the particular state that it is. This fact supports the view that mental states are individually owned and are identifiable as those of a particular, individual person.

The consideration of memory theories and memory research further supports the claims of this chapter. As memory theories have evolved, they have become more aware of memory's holistic structure, and of the fact that memories do not occur in isolation from other memories, or other mental items. Research into memory has shown that memory operates in a dynamic and reconstructive manner, and is informed by the original event, its storage conditions, and its recall conditions. This applies whether memory operation is normal or aberrant. These findings do not support the atomistic characterisation of memory. They show that merely an appeal to particular memories or small sets of memories cannot can assist in the elucidation of personal identity, since this merely shifts the problem of the unity of the person to the unity of mental states.

Indeed, the above findings support psychological holism. They support the claim that memories are not isolated mental events, but are part of a network of similar events. In view of these findings, it difficult to conceive how a discrete or impersonal account of memory could be defended. The integrally related interplay of factors which characterises memory experiences indicates that memories are predominantly holistically constructed, and consequently, are unique to their owners, and to the times and occasions at which they occur. In addition to its inability to account for the relations between mental states, the psychological continuity also has difficulty in accounting for the relation between the mind and the body. This topic is dealt with in the next chapter.

Chapter 5 Indiscrete Bodies

5.1 Introduction

It should already be evident from the discussion in the previous two chapters that the psychological continuity criterion is inadequate to account for certain aspects of mental content, such as how mental content relates to external objects, and how mental content maintains its internal structure. Further investigation shows that the psychological continuity criterion is similarly inadequate to account for the mind-body relation. According to the psychological continuity criterion, although persons typically have bodies, particular bodies are incidental to particular sets of mental states. Whether a person has one body or a completely different body is irrelevant to a person's mental life. It is taken to follow from this that because mental life is tied to personal identity, particular bodies have no relevance to personal identity either. One consequence of this view is that although mental life and personal identity are always accompanied by the body, they are not tied to any *particular* body. This conception of the mind-body relation can be seen from the psychological continuity criterion's view that psychological continuity arises essentially from brains or brain parts, to the exclusion of other body parts. The restriction of personal identity to the brain arises from the view that personal identity is grounded in causal connectedness between mental states, regardless of the type of cause which holds or the type of mental content involved. According to Parfit's thought experiments, the supposed separation of brains and bodies is taken to show that psychological continuity is maintained in the brain rather than in the body, and consequently, that personal identity is also maintained in the brain rather than in the body. In this chapter, I will argue that this view misconceives the relation between the body and psychological continuity, and consequently, between the body and personal identity.

Excluding the body as a condition of personal identity is consistent with the psychological continuity criterion's reductionist, atomistic conception of personal identity. This conception treats the elements that compose personal identity as if these elements were discrete and independent of each other. Minds and brains are treated as if they were independent the body. It is assumed that a person's mind would be substantially unaltered if the person's brain was removed from the body of which it is part, and then placed into a different body. Thought experiments which discuss brain bisections and brain transfers treat these brains and brain parts atomisticly, as if they would operate in precisely the same way after the transfers had occurred. According to this conception of the mind-body relation, the body itself has no input to personal identity, and therefore, when changes in a person's embodiment are supposed to occur, personal identity is taken to be unaltered.

In contrast to this atomistic conception of the body and personal identity, I present a holistic view of the body's relation to personal identity, which, while acknowledging the connection between the mind and the brain, also acknowledges the connection between the mind and the body. The holistic views conceives of the brain as an integral part of the body, rather as some sort of 'add-on extra' which can be arbitrarily removed or swapped. This view recognises the actual relation which exists between the brain and the body, rather than postulating one which is artificial and contrary to reality. Brains are not objects discrete from the body, but are an important *part* of the body, such that they are affected by the particular body of which they are part. This is an important point in this debate, as it is assumed that putting brains in different bodies would bring no change to mental life; however, when it is realised that mental life is a product of the brain-body unit, rather than just of the brain, it becomes evident that this is not the case. Many bodily functions affect mental life, as can be seen when the brain's relation to the rest of the body is better appreciated. These points, which will be

developed in this chapter, make it evident that it is not merely the *qualitative* identity of one's body that is relevant to personal identity, but that it is the *numerical* identity of one's body which is at issue. Because, as will be shown, the development and functioning of the mind is intimately tied up with the development and functioning of the body, the separation of them, suggested by psychological continuity theorists, actually makes no sense.

Brains are integrally connected to all parts of the body on a number of levels, ranging from the minute cellular level, to the level of the body's major systems. There are eleven of these systems, and the brain is crucial to the operation of each of them.[114] The brain itself is part of the body's nervous system. The nervous system is the body's major means of communication and control, and is extensive and complex. Its two main subdivisions are the central nervous system and the peripheral nervous system. Major components of the central nervous system include the brain and spinal cord, while those of the peripheral nervous system are the afferent system and the efferent system, comprised respectively of afferent and efferent neurones. The efferent system conveys impulses to the central nervous system from the body's extremities, while the afferent system conveys impulses in the reverse direction (Tortora 1980), pp 362-433. The main purposes of the nervous system are to maintain (with the endocrine system) homeostasis,[115] and to stimulate movement. This is effected by sensing and interpreting changes in the body, and initiating the appropriate action (Tortora 1980), pp 362-368.

---

[114]The body has eleven major systems, whose operations jointly keep the body alive and functioning. While each system operates independently, each system's operation is crucially influenced by the operations of the others. These systems are: the *Integumentary, Muscular, Skeletal, Digestive, Cardiovascular, Lymphatic, Respiratory, Endocrine, Nervous, Urinary,* and *Reproductive* systems (Tortora 1980), pp 80-86; 214-242; 92-94.

[115]'Homeostasis' refers to the state of stability within an organism.

Through its operation within the nervous system, the brain is connected to, and influenced by the body's other systems. For example, the circulatory system runs throughout the whole body, including the brain. Nutrients, drugs, poisons, or any other substances which circulate through the body, also circulate through the brain, and affect the brain accordingly. Similarly, the body's other systems affect, influence, and contribute to the life and operation of the brain, and the brain likewise contributes to the life and operation of these other systems, and through them contributes to and maintains life in all parts of the body. This means that the operation of the brain and the extent of its capacities, is intimately and directly tied to the body of which the brain is part. It follows from this that the mental life which we associate with the brain is actually a product of the whole body, rather than just of the brain.

Understanding this connection between the body and mental life is crucial to the personal identity debate, because it makes evident the important point that the *particular* body which one has is a factor in the type of mental life that one has. It is not just the case that the body is connected to mental life and that's that. It is the more particular point that which body one has *makes a difference* to the mental life that one has. Understanding this integral connection between the body and mental life makes it evident that because mental life is so intimately tied to personal identity, the body is inevitably tied to personal identity also. This conception of personal identity conceives of bodily identity as an essential feature of personal identity, rather than one which is merely incidental to it

Normal recognition procedures support the view that bodily identity is commonly taken as a major determinant of personal identity. In the normal course of events, it is by means of physical characteristics that we identify each other. When we encounter each other in our homes or work-places, we recognise these physical characteristics because they are accessible to us and because we are familiar

with them. They are the natural means by which we know and recognise each other. In situations where these normal recognition procedures may be unavailable, or where they may need checking, the body provides the unique identity-determining characteristics of DNA and fingerprints.[116] These features are standardly accepted as being identity-determining, as being unique to each person, and as not being duplicable (Norton and Esposito 1994), p 141; P 781.[117]

Physical characteristics are also important to self-recognition. Knowing who we are, recognising ourselves *as* ourselves, and knowing how we feel about ourselves is very much tied to our bodies.[118] We have an instinctive sense of ourselves as being persons with particular bodies, and not just as having *any* body. Imagine that if you could never be sure that you would wake up with the same body, or that you had no sense of what your body was like, what it looked like, or what it felt like, or how big or small it was, would your sense of yourself be the same as it is now? It seems to me that without the sense of our bodies as *particular* bodies, our sense of ourselves would be very much depleted, and somewhat strange.

The psychological continuity criterion does not recognise this integral connection between the body and personal identity. It assumes that the brain and body swapping that occurs in thought experiments can tell us something about personal identity in real life. But the brain and body swapping which goes on in thought experiments does not occur in real life. There is no reason to assume that the kind of identity changes described in these scenarios would occur in the way

---

[116]'DNA' is the organic chemical whose complex molecular structure codes genetic information for the transmission of inherited traits. 'DNA' refers to Deoxyribonucleic acid. Its 'unique structure confers great stability, allowing it to act as a template for the production, or replication, of new DNA molecules or of a related molecule, RNA (ribonucleic acid) which mediates the synthesis of proteins by the cell' (Norton and Esposito 1994), p 140. A more precise description is difficult here as it is complex and lengthy. Sufficient for our purposes here is that it is a *unique* genetic marker. Fingerprints are impressions made by the ridges on the ends of fingers.

[117]DNA and fingerprints are unique even to identical twins.

[118]Bodily self-awareness and self-knowledge is dealt with in more detail in the next chapter.

supposed. But what we can know is that the whole idea of personal identity disconnected from bodily identity is totally impracticable and unworkable. If we could not, as a general principle, rely on physical features to identify each other, personal, family, social, and other relations would be totally disrupted and unmanageable.[119] Think about it. If you could not rely on physical features as a means of person identification, you could never know for sure that the persons with whom you interacted in the normal course of events really were the persons you took them to be. Without bodily identity, you would be supposed to rely on the amount of 'psychological connectedness' in persons' minds to know who these persons were. But it is a mystery how you could discover what this amount was. How could you measure the amount of psychological connectedness present in a person's mind, and how could you know if was the right amount? The problem here, then, is not just that we should change our recognition procedures, but rather that we *could not* change these procedures. Without bodily recognition as a prime determiner of personal identity, we could never be sure who we were with, or who we even were ourselves. Physical features are, in most normal circumstances, absolutely essential to our interactions with each other, to our identification of each other, and to our identification of ourselves.

The body's role in mental life also supports the view that bodily identity is intimately tied to personal identity. As shown in Chapter 3, minds are not internally constituted, but are constructed in relation to certain external features. Recall that these features relate to our *Umwelt,* and concern the 'functional cycle,' in which the immediate environment, survival needs, perceptual apparatus, and activity are understood as a 'systematic whole,' that is, as an interrelated and mutually supportive set of factors, in which an organism's perceptual response initiates its

[119]Of course, mistakes in identification can be, and often are, made. But these mistakes are made against a background in which most identifications we make in the course of our lives are correct. Making a few mistakes is not the same thing as living in a world in which one could not, in general, rely on bodily identification, especially in the case of close relatives and friends. In spite of its lack of total certainty in all instances, unless we could rely on bodily identification, normal life and relationships would be unworkable.

activity, and in which its activity influences and defines its perceptual response.[120] For humans, the functional cycle is tied to both the body and the mind. The particular body one has determines one's survival needs, perceptual responses, activities and so on, such that were one to have a different body, these things would be different also. And our particular perceptual responses form the basis of our thoughts, such that if our perceptual responses were different, our thoughts would also be different. Through our perceptual responses there is thus an integral relation between the particular body one has and the particular mind that one has. And because personal identity is intimately tied to particular minds, it is at the same time intimately tied to particular bodies. The relevance of individual bodies to individual sets of mental states will become more apparent as this chapter progresses.

The claim that the body is irrelevant to personal identity is not an accidental feature of the psychological continuity criterion. If psychological continuity was recognised as a product of the whole body, rather than just of the brain, then the characterisation of psychological continuity independently of bodily continuity could not be made, and thus, personal identity could not be characterised solely in terms of psychological states, to the exclusion of bodily states. In other words, the arbitrary separation of mental states and bodily states could not be proposed if the holistic structure of the body was recognised and if the integral relation between mental states and the whole body, rather than just the brain, was taken into account. Due to this integral relation between body, mental life, and personal identity, I will argue that the neglect of the body in the account of personal identity by the psychological continuity criterion produces misleading conclusions about the relation between personal identity and the body. I will also argue that due to its influence on mental life, the body is decisively involved in psychological continuity, and correspondingly, also in personal identity. To further argue this case, I will first

[120]See Chapter 3.5.

examine some thought experiments in which brains and bodies become separated from each other. I will then consider some examples of brain-body interaction, and ways in which this interaction impacts on the relation between bodily identity and personal identity.

5.2 Irrelevant Bodies

The psychological continuity criterion's neglect of the body begins with the criterion's particular formulation. By characterising personal identity as holding in 'overlapping chains of psychological connectedness,' to the exclusion of bodily connectedness, the relevance of embodiment to personal identity is overlooked. While this does not mean that persons are taken to be disembodied entities, it does mean that the body's role in personal identity is not taken into account. This neglect of the body implies that the body is not essential to the identification or reidentification of persons, or to the realisation of psychological continuity. Parfit's representation of the relation between bodies, brains and minds is particularly problematic in this regard. Bodies are treated as silent partners in the realisation of psychological states. Brains are treated as discrete and malleable items, capable of dissection, manipulation, and re-embodiment, without any significant disruption to a person's continuing psychology. This implies that neither the identity of the body, or of its parts, is significant to personal identity.

The rift between bodily identity and personal identity mirrors a division within the Lockean characterisation of personal identity from which the psychological continuity criterion descends. Locke holds that bodily continuity consists in the continuity of organic matter, whereas personal continuity consists in the continuity of rationality and consciousness. Locke distinguishes between substance, body, and person:

> It is not therefore unity of substance that comprehends all sorts of identity, or will determine it in every case; but to conceive and judge of it aright, we must consider what idea the word it is applied to stands for: it being one thing to be the same

> *substance,* another the same *man,* and a third the same *person, if person, man,* and *substance,* are three names standing for three different ideas;- for such as is the idea belonging to that name, such must be the identity; which, if it had been a little more carefully attended to, would possibly have prevented a great deal of that confusion which often occurs about this matter, with no small seeming difficulties, especially concerning *personal* identity, which therefore we shall in the next place a little consider (Locke 1959), 2.27.8.

Persons and bodies normally occur together, but according to Locke, may not always do so. Locke cites the case of a 'rational parrot' which was able to converse as competently as any human person. Locke then proposes that persons are essentially rational, *psychological* beings (Locke 1959), 2.27.9-11. Because human psychological characteristics are not held to depend on particular bodies, it is assumed that the body is incidental to personal identity. Subsequent psychological continuity theories, such as Parfit's, inherit this characteristic by disregarding the body in personal identity accounts, and consequently, denying that bodily identity is relevant to personal identity.

5.3 Disembodied Memories

The psychological continuity criterion's disdain for the body is evident in its various thought experiments, in which brains are dissected and transferred to different bodies. In these thought experiments, the quality of a mental experience associated with a particular brain part is taken to be unaltered, even though the brain part concerned is removed from its original location, and placed into the skull of a different person. The earlier thought experiment concerning Jane and Paul is an example of this type of scenario.[121] Recall that this thought experiment involves the transfer of a small part of Paul's brain into the skull of Jane. Subsequent to this transfer, Jane is supposedly able to experience Paul's memories. Jane is supposed to remember events such as walking across marble paving, hearing bird-cries, and shaving. It is taken that a person with one body is

[121]The Jane and Paul thought experiment is referred to in Chapter 4.3.

able to 'remember' events experienced by a person with a different body. According to this scenario, although these memories are first-person memories, they are unaffected by their different embodiment. This view implies that although the body is involved in experience, it plays no specific role in experience. If the body *was* recognised as a specific factor of a person's experiences, it would be acknowledged when considering the person's ability to remember such experiences. The relevance of particular bodies to experiences and memories becomes clearer when we reconsider the Jane-Paul case.

In this scenario, we are told that Jane has first-person memories of events which were experienced by Paul. First-person memories are memories in which one remembers one's *own* self performing an action, rather than remembering someone else performing it. These actions are remembered from a first-perspective on the world, which means, not only from the perspective of one's own mind, but also, from the perspective of one's own body. It is difficult, as we saw earlier, to conceive how one person could have first-person memories of events experienced by a different person. For example, when I remember myself walking, my memory has a first-person quality, which means that I remember it from the point of view of *my body.* It is specifically from the standpoint of my body that I see, feel, hear, and interact with the world. There is thus something very odd in the idea the Jane could have first-person memories of events experienced by Paul, when she does not have Paul's body. We need to consider more closely what this could mean.

The transferred memories consist of various physical activities which were performed by Paul, but remembered by Jane. As Jane did not perform the activities in question, her ability to remember doing so, however, cannot be taken for granted. For example, Jane is supposed to remember walking across marble paving. But if Jane was physically different from Paul, say much shorter, or much

heavier, it is difficult to see how her memory could be from Paul's perspective, or could have the same quality as Paul's memory. This problem becomes more obvious if we consider how Jane might remember walking if she was a paraplegic, or born without legs. Having never walked, it is difficult to know how she would be capable of 'remembering' walking. Her legs and other body parts would never have been engaged in the way required. Of course she could imagine what it felt like, but this would hardly qualify as a memory, and there is no guarantee that her imagination could match the reality.

Jane is also supposed to remember the cries of gulls in exactly the same way as Paul had heard them, and as he would have remembered them. Again, we might wonder how Jane could do this if she was born deaf, and had never heard sound before. How would she even know what 'sound' was, or be able to know that the 'sound' in her memory was the cries of gulls? And finally, the shaving episode is intriguing, as it is difficult to see how Jane could remember shaving, when she does not have the face that was shaved, or even a face which has ever been shaved. The memory is supposed to be a first person memory, but it is difficult to see how a memory of Paul's chin could be a first person memory, or how a memory of shaving Jane's own chin could qualify as 'Paul's memory.' If Jane and Paul's bodies were similar, and if their habits and behaviours were very much alike, it is conceivable that their memories could have some similarity. But if there were significant differences between their bodies or their behaviours, it is difficult to conceive how their memories could be alike, or, in some instances, even recognisable. Notice here also that the role of the body in memory means that the change in bodily identity renders questionable the continuity in identity of memory.

The problem with this thought experiment is that it fails to recognise the role of the body in mental life. In failing to recognise this role, it misses two important points concerning experience. The first concerns the first-person perspective of

experience. The failure to include the body in the account of mental states leaves the first-person perspective inadequately defined. A more adequate appreciation of the first-person perspective indicates that it makes no sense to claim that one person could remember the experiences of a different person. The second point concerns bodily differences. The thought experiment leaves no place to account for bodily differences, and for the fact that particular bodies are tied to particular sets of mental states, such were a person's body to be different, that person's mental states would be different also. This indicates that bodily identity is indeed a fact in the identity of mental states. These two points are explained in more detail.

The first-person perspective of an experience is the perspective of the person who has the experience, or, in other words, the experience from the point of view of the experiencing subject. Coinciding with the first-person perspective is the sense of ownership. As explained earlier, this sense of ownership does not come separately from the experience itself.[122] Knowing that an experience is mine is synonymous with knowing that the experience occurs from my point of view. This point of view coincides with the location activities, structure, capacities, and so on, of my particular body, as opposed to those of a different body. Thus, having a first-person perspective of an experience coincides with knowing that I engaged in that experience with *my* body.

If experience requires a point of view, it follows that the memories of experiences also require a point of view. When I remember an incident, I do so from the perspective from which I experienced that incident. It would make no sense to say that I remembered it from the point of view of your experience, rather than from the point of view of my own experience. Thus, just as experiences are in part defined by viewpoints and ownership, so also are the memories of experiences. It follows from this that if bodies are implicated in the viewpoint of experience, they

[122]This point was discussed earlier, see Chapter 4.4.

are also implicated in the viewpoints of memories. Just as I experience an incident from the viewpoint of my body, I also experience the memory of that incident from the viewpoint of my body. In view of this, there is something very odd in the idea that I could remember the experiences of a person with a different body, as if these experiences were my own. What exactly does this mean? Does it mean that I mistakenly remember myself doing something, which in fact I did not do? Or does it mean that I remember someone else doing something? If it is the first, it is a mistaken memory, a delusion, false belief, hallucination, or similar, but by definition, is not a genuine memory. If it is the second, it is not by definition a first-person memory of myself performing an action, but of someone else performing that action. It simply is not possible to have it both ways. I cannot both know that an experience was had by someone other than myself, and simultaneously claim to remember myself having the experience. How can I both admit to not having had an experience, and at the same time as claim to remember having had it? This is incoherent, and makes no sense whatsoever.

The second problem with this thought experiment is that in failing to account for the role of any *particular* body, it fails to recognise that because all bodies are different, their contribution to mental life is also different.[123] Each body has its own set of experiences. These experiences are defined by various elements, such as the composition, structure, activities, location, relationships, and so on of the particular body concerned. Correspondingly, the mental life of each person is tied to, and indeed, constructed by these same elements. Thus, the set of elements which comprise the experiences, and hence the mental life, of a particular individual person are not the same as those which comprise the experiences and mental life of a different individual person. It makes no sense, therefore, to assume that the mental life of one individual would be remain unaltered if that

[123]This point is not overturned by the claim that identical twins have qualitatively identical bodies. Right from birth, the perspective and interactions with the world of one twin is distinct and separate from that of the other. These differences become manifest in the body of each twin, such that the composition, structure, and appearance of each twin continues throughout life to differ from that of the other.

individual were to have a different body. In summary, the theoretical separation of brain and body, exemplified in this thought experiment, incurs metaphysical consequences concerning the relation between experience and embodiment which cannot be sustained.

The incoherence of brain-body separation is supported by Marya Schechtman. Schechtman recognises that the interaction that occurs between the body and the brain is an important factor in any experience. Schechtman claims that understanding any mental functioning includes knowing what occurs in the body as well as in the brain. Both physiology and psychology, not to mention common sense, demonstrate that:

> it is virtually never the brain alone which is active in mental life. The brain is in constant interaction with the rest of the body, and the physiological understanding of mental functions almost inevitably involves not only an understanding of what is happening in the brain, but in the entire feedback loop between the brain and peripheral systems (Schechtman 1997), p 152.

In cases where injuries supposedly leave the mind unaffected but the body damaged, such as when vision or hearing are impaired, experience is drastically affected and mental functioning is also altered accordingly. Schechtman argues that rather than regarding mental activity as confined to the brain, we should recognise it as a product of the entire body and of factors beyond the body, such as the immediate environment. She refers to this as the 'distributed view.' The distributed view recognises that the body and the environment are implicated in the mind, as opposed to the 'standard' view, which holds that the mind is located entirely in the brain. Schechtman claims that in denying recognition of the brain's external factors, the standard view presupposes a self 'inherited from Cartesian dualism - a psychological subject living inside of a body' (Schechtman 1997), pp 151-153. Schechtman argues that in neglecting the body's relevance to personal

identity, much contemporary literature on the topic is based on little more than speculation:

> Virtually the whole of the contemporary literature on personal identity turns around puzzle cases in which the brain is manipulated in some way or another - if not transplanted then bisected, or altered bit by bit. It is absolutely crucial to these discussions that the brain be considered the locus of the psychological subject. This means that the whole of the modern personal identity literature rests on a questionable set of assumptions (Schechtman 1997), p 160.

That the psychological continuity criterion rests on such a weak foundation is often missed by personal identity theorists. But as shown above, in neglecting to acknowledge the body's influence on mental states, the psychological continuity criterion overlooks an important influential and constraining factor of mental life, and consequently, and important factor in the realisation of personal identity.

5.4 The Unity of Consciousness

Another kind of thought experiment which involves discontinuity between the body and the brain is based on brain bisection surgery. Brain bisection is a surgical procedure used to control epilepsy, in which the fibres between the right and the left brain hemispheres are severed (commissurotomy). Following one cited example of this surgery, a person was claimed to draw a pipe with one hand, while writing the word 'pencil' with the other. In another example, a person was claimed to have pushed his wife away with the left hand, while his right hand embraced her. Parfit claims that these and similar responses indicate that two separate streams of consciousness are present in a single-bodied individual (Parfit 1984), pp 245-246. If each stream of consciousness is taken to represent a separate, individual person, this would mean that two persons were present in the one body. If two persons could be present in a single body, this would support the case that bodily identity is not tied to personal identity. Parfit's extrapolates from these examples to further this case.

Based on the above argument, Parfit imagines that he is fitted with a device which prevents the two hemispheres of his brain from communicating with each other. The operation of this device supposedly allows him to divide his mind into two separate streams of consciousness. With each stream unaware of the other, he could work separately on exam problems, reuniting his mind at will. Parfit claims that each stream would experience a unified consciousness, as this is what happens in actual cases:

> And it is a fact that people with disconnected hemispheres have two separate streams of consciousness - two series of thoughts and experiences, in having each of which they are unaware of having the other. Each of these two streams separately displays unity of consciousness (Parfit 1984), p 247.

Expanding his example, Parfit imagines that it is possible to divide a brain many times, resulting in many branches of consciousness emanating from a single, individual person. Each brain part is placed into a different body, allowing one 'person' to become many 'persons.' The survival of these 'persons' is unaffected by their different embodiments. But because non-branching psychological continuity is violated, personal identity is overturned, and, as a result, becomes meaningless. Parfit argues that because personal identity loses value in this way, having more concern for oneself than one has for persons other than oneself is irrational. He also argues that even if he were himself to have surgery which would destroy his own psychological continuity, and his life would become 'worse than nothing,' he could not justify egoistically based concern. Because his belief that the resulting person would still be him is unjustified, so also is the precedence of self-concern over other-concern. Parfit argues that whether the future person is him, or whether it is someone else makes no difference as to whether he should care about what happens. He holds that neither split psychology, nor altered bodily continuity has any relevant impact on personal identity. These factors are not significant when speculating on 'what happens' when persons 'divide.' If asked why the brain was singled out from other body parts, Parfit would reply:

> Because the brain is the carrier of psychological continuity, or Relation R, if this is why the brain is singled out, the continuity of the brain would not matter when it was *not* the carrier or Relation R. The continuity of the brain would here be no more important than the continuity of any other part of the body. And the continuity of these other parts does not matter at all Parfit 1984), p 284.

Parfit claims that physical continuity is the least important element in a person's continued existence, as it is outweighed by features such as relations, achievements, and other psychological factors. He claims that the body is irrelevant to personal identity, and is, therefore, also irrelevant to these concerns (Parfit 1984), pp 245-284. If, however, personal identity is ultimately meaningless, it is unclear how an individual's relations and achievements could have any meaning, or indeed, to whom or what they could mean anything to.

A major problem with this thought experiment is that the foundation on which it is based is faulty. Parfit bases much of his argument on the view that commissurotomies produce two streams of consciousness in a single-bodied individual. This claim is not only incorrect, but is also incoherent. While it is true that some researchers initially proposed that two fields of consciousness may result from brain bisections, this view has since been modified.[124] Gazzaniga is an example of one researcher whose views have changed in this regard. Early researchers in this field became aware that the co-ordinated functioning of the right and left hemispheres becomes disrupted following brain bisection surgery. Gazzaniga initially considered that this disruption means that two distinct conscious mental spheres have been produced:

> With it sectioned, the two halves become two different conscious mental spheres, each with its own experienced base and control system for behavioral operations. Just as conjoined twins are two different people sharing a common body, the callosum-sectioned human has two separate conscious spheres sharing a common brain stem, head, and body (Gazzaniga 1970), p 1.

[124]For a report on this original surgery, see: (Gazzaniga 1970).

Subsequent to more considered research, however, Gazzaniga and others deny that altered mental functioning means that consciousness is actually split. Rather, these researchers recognise that following brain bisection surgery, the processing of information between the two hemispheres of the brain is disrupted, but that this does not imply that two separate streams of consciousness are present. In one example in which objects were flashed in the left visual field, but in which the presence of objects was verbally denied, researchers conclude that the ability to manually retrieve these objects means that both hemispheres *were* aware of the objects' presence, even though the subject concerned could not verbalise this awareness:

> clearly, the right half-brain knew the answer, because it reacted appropriately to the correct stimulus. That each half-brain could process information outside the realm of awareness of the other raised the intriguing possibility that the mechanisms of consciousness were doubly represented following brain bisection. The implications of this controversial possibility were far-reaching and attracted the interest of philosophers and scientists alike. However, while the conscious properties of the talking hemisphere were apparent, the view that the mute hemisphere was also deserving of conscious status was widely criticized and generally rejected. Consequently, subsequent studies focused on elucidating the nature of information processing in the right hemisphere (Gazzaniga and Le Doux 1978), p 5.

Gazzaniga claims that some early responses to brain bisections were misunderstood, leading to 'overdramatized accounts of the uniqueness of mind left and mind right', and to 'pop versions of hemisphere function[ing]' (Gazzaniga and Le Doux 1978), p 6. He also claims that in many cases, these errors were made by persons who had no first-hand knowledge of the patients involved.

When considered in more detail, the notion of 'split consciousness' itself is inherently problematic. It is difficult to even conceive of what it could mean. For example, how could a person ever know that she had two separate 'fields' of consciousness?' If the fields were separate, it would not be possible, when

experiencing one field, to know that the other one existed. If you did know such a thing, the 'other field' would not be separate, but would be part of the one you were currently experiencing. While Parfit's argument supposedly addresses the issue of the first-person viewpoint, in fact, it does not do so. A first-person viewpoint corresponds with, and is defined by, a single stream of consciousness, from the perspective of a particular body. It is thus incoherent to propose that there could be two streams of consciousness, and thus, two different first-person viewpoints, emanating from one, single body.[125]

More recent studies, such as those of Grant Gillett, also question Parfit's arguments concerning brain bisections.[126] Gillett argues that research does not support that brain bisection surgery occurs in the way that Parfit describes, nor that it produces the results that Parfit claims. According to Gillett, the epilepsy surgery on which Parfit bases his argument involves the information-processing parts of the forebrain known as the diencephalon and the cerebral cortex. Surgery does not usually mean that all fibres between the two hemispheres are severed. Following surgery, diverse responses from each hemisphere to stimuli sometimes occur. This, however, is usually in controlled or manipulated situations, where each hemisphere is fed with different material, perhaps by information being flashed quickly across a screen, such as in the case described by Gazzaniga above. While these test situations may produce different responses from each hemisphere, normal life is usually little affected. Situations described by Parfit (such one hand pushing and one hand embracing), are the exception rather than the rule. Gillett claims that the disruption of information processing is usually only

---

[125]An objector could reply: 'Well, what about multiple personalities? Are they not evidence that more than a single stream of consciousness is possible in a single-bodied individual?' This is not necessarily the case. MPD (Multiple Personality Disorder) is by no means a proven case that multiple consciousness in a single-bodied individual is possible. Many theorists see MPD as a single consciousness which is disordered and dysfunctional. Sufferers of the disorder do not have completely separate mental fields, much knowledge, especially bodily knowledge, is shared among so- called alter personalities. See (Kennett and Matthews 2001).

[126]Grant Gillett is not only a philosopher, but also a brain surgeon who has first-hand knowledge of the issues discussed here.

temporary. The real question is whether there is one subject of experience, rather than whether there are two streams of consciousness (Gillett 1986), pp 224-226.

Gillett notes that patients are aware, not only of making mistakes when giving diverse responses, but also that it is *they* who are making them, not someone else:

> The fact is that the person realises that *he* has made a mistake, not that someone, perhaps contingently related to him has made a mistake which he has the knowledge to correct. As far as he is concerned, the person in error and the person who is not are he, himself, one person and one mind, but he is not functioning properly (Gillett 1986), p 226.

Gillett argues that too much is read into the fact that each hemisphere responds to the same stimuli in different ways. This, of itself, does not indicate that there is not one single experiencing subject. In these cases, what happens in the brain is misinterpreted. By their nature brain parts divide labour, so it is no mystery that the communication between different brain parts is lost when these parts are severed. That one brain half not trained in language cannot respond linguistically is exactly what *should* be expected. As new connections form within the two brain halves, many apparent anomalies decrease over time. If we are to fully understand thought, we need to appreciate what is *necessary* for thought. Gillett notes that this requires an appropriately constituted subject for whom thoughts are possible. Such a subject has a set of beliefs, a conceptual framework, an environment, and a set of relations within which to situate and identify herself. Having a single mind does not mean having a perfectly integrated mind. To struggle with integration is not unusual. For these brain-disturbed patients the struggle is simply more acute:

> Because they do try to reintegrate their information, or make best use of their disrupted brain function in tackling the tasks they are set, they can properly be said to be struggling with certain confusions to which they find themselves subject rather than

> to have become two mutually independent streams of consciousness which are in no more than a contingent relation to each other (Gillett 1986), p 227.

According to Gillett, the evidence strongly supports the idea that a single person, with a single sense of identity, remains following the severing of brain fibres. Analysing Parfit's imagined proliferation of brain divisions, Gillett argues that personal identity involves relationships, reciprocity, a history, a biology, all of which are interrelated and which could not be 'replicated' in some arbitrary way. He claims that persons are entities with a certain type of interconnected description, and that it is not clear that such a description could be given to Parfit's reduplicated entities. These entities would be severely impaired, and on current biology and psychology, would be unlikely to have any kind of meaningful mental life at all (Gillett 1986), pp 227-229. In summary, the above criticisms indicate that the foundation for Parfit's break between bodily identity and brain identity is not well-founded, and consequently, that bodily identity and personal identity may not be as discrete as Parfit implies.

5.5 Brain-Body Integrity

The foregoing arguments imply that bodies, brains, and minds may not be as autonomous as the psychological continuity criterion leads us to believe. The view that these things are fundamentally connected is supported by consideration of ways in which the precise functioning of the brain is affected by, and indeed tied to the functioning of the body to which that brain is connected, or more precisely, of which it is part. When the connection between the brain and the body is fully appreciated, it becomes apparent that the mental life of a particular, individual person is, at least in part, the outcome of the interactions and experiences of a particular brain-body unit, rather than of a particular discrete brain. This does not mean that a person's mental life is reducible to the body, but only that it is constrained and influenced by the body, and consequently, that bodily identity is

an important component of personal identity. In other words, even the identity of a full brain is insufficient on its own to capture personal identity.

As mentioned earlier, the integrity of the brain-body unit is demonstrated by the way in which the brain is related to the body through the body's various systems, key examples of which are the cardiovascular system and the nervous system.[127] The cardiovascular system is responsible for the circulation of the blood throughout the whole body. This system ensures that blood, which circulates through the body tissues, also circulates through the brain tissues. Blood contents which nourish or damage the body, also nourish or damage the brain. Indeed, the very maintenance of the brain depends on the nutrients which it receives from the blood-supply. The blood-supply receives these nutrients from ingested food. Where the body's nutrition is deficient, brain functioning is impaired accordingly, and mental life is affected. Even the amount of water that one drinks affects the brain, as water contributes to the brain's correct functioning.

Brain activity is also affected by substances such as alcohol, coffee, or cigarettes. These substances alter the chemical activity which occurs at the brain's chemical synapses, causing stimulation to parts of the brain. Moods and behaviours are altered accordingly. Substances that are released into the bloodstream by the brain, such as hormones (under the control of the hypothalamus via the endocrine system) also affect the relevant body parts, examples being the onset of puberty or menopause. These examples demonstrate that what happens in the brain is entirely contingent on what happens in the circulatory system. Because brain activity and bodily activity are intimately related in this way, it cannot be taken for granted that, were a change in embodiment to occur, a person's mental life would not be significantly affected (Curtis 1983), p 825; (Tortora 1980), pp 306-330; 471-509; (Damasio 1994), p 88. And because mental life is so significantly related to

[127]The integrity of the brain-body unit was referred to in Section 5.1 of this chapter, and is now explained in a little more detail.

personal identity, it cannot be taken for granted that a change in embodiment would not affect personal identity also.

Another system which significantly influences brain activity is the nervous system. This system ensures that signals from all body parts are relayed to the brain, and that the brain correspondingly is in contact with all body parts. All activities, whether they involve joints, muscles, or organs, or whether they involve the regulation of homeostasis, are affected by this system (Tortora 1980), pp 362-433; (Damasio 1994), p 88. The interrelatedness between the body and brain cannot be overestimated. The use of separate linguistic terms 'brain' and 'body' is misleading, as it gives the impression that the items themselves to which these terms refer are disassociated from each other. While this may be the case in a dismembered corpse, it is not the case in a living, human organism. Antonio Damasio claims that even the description of the body and brain as an 'indissociable organism' is an oversimplification. As an indissociable organism, the interaction between brain and body involves constant information relays between each other's parts. It is only due to this interaction that any part of the body or brain can operate in the way that it does, or indeed, at all (Damasio 1994), p 88.

A key element of brain-body interaction is the transmission of signals between the brain and the rest of the body, and between the parts within the brain. These signals control all brain and body functions. They are passed between the end-points of neurones in the brain and other parts of the nervous system called *synapses.* Synapses emit chemical signals to other synapses. The synapses do not actually touch, but are connected only by signals. Electrical signals pass directly from one synapse to another. Chemical signals are released across the *synaptic cleft,* in the form of *neurotransmitters.* Neurotransmitters are chemical agents released by neurones, which affect either other neurones, or glands or muscles, by altering the 'electrical state or activity' (Curtis 1983), p 1101. This

alteration is in the form of either excitation or inhibition. The chemical signals travel across intercellular fluid. On reaching their destination, neurotransmitters combine with postsynaptic cells, after which the neurotransmitters are destroyed, or diffused. Prominent neurotransmitters in the peripheral nervous system are acetylcholine and noradrenaline (Curtis 1983), pp 773-775.

Noradrenaline is also produced in the hypothalamus, and in other parts of the limbic system, and is thought to be involved in arousal and attention levels. Low levels of noradrenaline at certain synapses has been associated with severe depression, a condition which affects mental functioning. Medication to increase noradrenaline at these sites is sometimes used as a means of correcting this condition. Such cases provide examples of how what occurs in a particular body directly influences and constrains what happens in a particular brain and how mental life is correspondingly affected. Were the brain of a person who suffered from depression, and who had just taken medication, to be transferred to an unmedicated body, it is difficult to see how that person's activities or mental life could be the same as it would have been had the transfer not occurred (Curtis 1983), p 825; (Tortora 1980), pp 416-417. This example of the tie between the activity of the nervous system and mental activity further demonstrates the integral connection between mental life and bodily life. Moreover, because mental life is intimately connected to personal identity, this tie also demonstrates ways in which personal identity is integrally connected to the life of a particular body, rather than to the life of just *any* body.

Further evidence of the interrelatedness of the brain and the various bodily systems is apparent from recent studies of Alzheimer's disease. These studies demonstrate ways in which the brain, the circulatory system, and the nervous system are mutually related, thereby supporting the view that particular minds and identities are related to particular bodies. The (at the time of writing this thesis)

12-year old *Optima* project studies Alzheimer's and other aging diseases.[128] Alzheimer's disease incurs massive degeneration of nerve cells in the brain's cortex, resulting in severe cognitive defects, typically including memory loss in the elderly. The deficits are far more severe than, and not correlative with, the normal decline expected with age. Studies demonstrate that severe brain damage occurs over several years, leading eventually to extreme cognitive impairment, and ultimately, to death (Perry 1999), p 46. Although the study findings are incomplete, there is mounting evidence that environmental factors, including body chemistry, are crucially implicated in the disease.

Research in the 1960s found that the normal aging process involves microscopic changes in the brain.[129] Plaques of amyloid protein become deposited outside neurones in the brain, while neuronal fibres become 'tangled' inside the neurones. Mental life, particularly memory ability, is affected accordingly. While it was initially thought that the neuronal changes were early symptoms of Alzheimer's disease, it was later realised that this is not always the case. A task for Optima was to provide a biological 'state marker' by which to measure the course of the disease throughout a person's life. Comparisons could then be made between changes in the marker, and the changes of normal aging. A state marker was provided by use of CT scanning (X-ray computed tomography) which image-scanned the brain. In particular, the medial temporal lobe was the target, as post-mortem studies have found this to be the site with the most tangled neurofibres. Although this brain-part is only 2% of the brain's total volume, it includes the seat of memory, the hippocampus. Because the neurones in the temporal lobe link to other parts of the

---

[128]The Optima (Oxford Project to Investigate Memory and Aging) Project was set up in 1988 at Oxford University, England. Its aim is to study aging diseases, especially Alzheimer's disease. Alzheimer's disease is the most common cause of dementia in the elderly, presently affecting about four million persons in the United States. The percentage of persons affected doubles every five years above age 65 years. I am fortunate in receiving the information for this section from the Project's leader, Dr David Smith (Perry 1999), p 46.

[129]The research was carried out by Garyl Blessed, Bernard Tomlinson and Martin Roth (Smith 1998), p 98.

cerebral cortex, damage to the temporal lobe affects these other parts also. As evidenced by the CT scans, the medial temporal lobe of Alzheimer's sufferers[130] is much smaller than non-Alzheimer's age-matched controls. The amount of *atrophy* present in the medial temporal lobe equated with the *density* of neurofibrillary tangles present in the hippocampus following death. A relation is thus perceived between the neuronal tangles and the loss of synaptic connections, and hence the loss of tissue density in the brain (Smith 1998), pp 98-100.

Of the possible causes of degeneration considered by researchers, the most likely was that degeneration was triggered by a sudden brain event. Researchers reached this conclusion after examining the results of continuous brain-scans which were taken over several years of suspected Alzheimer's suffers. An incredible 15% per year medial temporal lobe shrinkage was detected in these scans, as opposed to the normal aging rate of 1.5% per year. Evidently, the cause of atrophy was catastrophic, and according to the researchers, clearly distinct from the causes associated with normal ageing. The direction of Optima's research then turned to considering possible factors associated with neuronal decay. Several likely candidates were produced, each of which could play a smaller or larger part in the decay process. According to researchers, the key factors relate to genes, body, and environment. Some of these factors could be: age, head injury with concussion, myocardial infarction in women, atherosclerosis, oestrogen deficiency in women, low education, and dietary factors. While these factors affect mental life through their effect on the brain, most of them are located outside the brain itself. They involve the body either directly, such as through hormone levels or diet, or indirectly, through environmental factors, such as education, age, and so on. These findings provide further examples of how factors associated with the body are crucially involved in brain functioning and mental life. And because brain functioning and mental life are crucially involved in personal

[130]That is, persons who eventually died of Alzheimer's disease (Smith 1998), p 99.

identity, the bodily factors which affect brain functioning and mental life are crucially involved in personal identity also. Indeed, when the involvement of the brain-body unity on mental life is considered in relation to diseases such as Alzheimer's disease, the connection between bodily identity and personal identity becomes even more apparent. For many crucial factors which affect memory loss (such as those mentioned above) are directly tied to the state of the body and to the environment in which the body is located, such that were these bodily states and environmental factors different, the person's mental life would be different also (Smith 1998), pp 101-103.

The continuing study of Alzheimer's disease further supports these findings. Epidemiological studies are still in process and final results are not yet available. But environmental and bodily influences are still held high as likely causal factors. Hormone and vitamin levels are being investigated. Some studies indicate that post-menopausal women who take HRT are at less risk than matched controls not taking HRT.[131] It is thought that oestrogen may be involved in protecting otherwise potential sufferers. Vitamin and dietary factors are also being examined. The researchers did not pay attention to these factors earlier in the study, but have recently realised their relevance. Low intake of vitamin C is thought to be a factor, and low blood levels of folic acid and B_{12} are also implicated. This conclusion comes from an interesting turn in research, in which the stored blood from deceased Alzheimer's sufferers was examined.

The examination of this blood revealed that the level of homocysteine, (a substance connected with stroke and heart disease) was 30% higher in the stored blood than in matched controls. Because the moderation of this substance is associated with levels of folate and B_{12}, more tests were done. It was found that the levels of these vitamins was down by 30%. Researchers hypothesised that in

some instances, there was a link between the abnormal processing of these vitamins, and the onset of Alzheimer's disease. The resulting pathology of this abnormal processing is similar to that of *transient ischaemia.* Transient ischaemia is a blocking of the arteries, a condition which results in there being reduced oxygen and an irregular glucose supply. Animal experiments which imitated the pathology of transient ischaemia produced biochemical and cellular changes similar to those of Alzheimer's. While there has not yet been the opportunity to fully test the hypothesis, researchers have records of four patients, who had memory loss in conjunction with folate and B_{12} deficiencies. Their memory problems disappeared following treatment with these vitamins. Researchers thus strongly suspect that factors such as diet are crucially implicated in the degenerative process which characterises this disease.[132]

The ability of diet and other factors associated with the body to affect the brain and mental life in this way provides compelling evidence of the *intrinsic* interrelatedness between the brain, the body, and mental life. We can see from the above research that mental life is not just a 'given,' but is entirely dependent on the functioning of particular brain-body units, rather than just on particular brains. This view of mental life does not support Parfit's characterisation of mental life as in some way discretely captured by the brain alone. Nor, as had been stated repeatedly throughout this chapter, does this view of mental life support the claim that personal identity can in any way be captured by psychological continuity to the exclusion of the particular, individual body with which that psychological continuity is associated.

---

[131]HRT= hormone replacement therapy. Often taken by before, during, and after menopause, with the intention of preventing the side effects associated with the reduced hormone production of menopause.

[132]When asked if he was excited about the dietary implications for Alzheimer's, as opposed to possible drug treatment, head of the Optima Project David Smith replied: '**A:** Yes, absolutely. It gives you a hope at long last of possibly being able to prevent some of this dreadful disease. That's why it's so exciting. I think therapy with drugs is a long shot and not too hopeful. But if you can come in before it starts and prevent it, you can save many millions of people from developing the disease. **Q:** Especially with dietary intervention? **A:** Yes, it's a relatively safe kind of thing that you can give' (Perry 1999), p 69. The researchers' conviction that diet influences brain structure and activity in this way supports the claim that brain and body are integrally related at a very fundamental level.

5.6 Summary

This chapter has addressed the question of whether bodily identity is implicated in personal identity. The psychological continuity criterion does not attach any weight to the body when determining what is crucial to personal identity. I have argued that this neglect of the body by the psychological continuity criterion produces misleading conclusions about personal identity. One of the most important of these is the view that the body is not a crucial factor in a person's psychological continuity, and that consequently, bodily identity is discrete from personal identity. According to this view, although persons are taken to be embodied entities, it makes no difference to a person's identity whether a person has one particular body or an entirely different body. Personal identity is thus in no way tied bodily identity.

In contrast to this view, I have argued that the body plays a decisive role in the realisation of psychological continuity, and that consequently, personal identity is, in part, defined by bodily identity. To defend these arguments, I have examined the psychological continuity criterion's atomistic conception of the brain, and its discrete characterisation of the brain-body relation. I have argued that thought experiments in which memories are supposedly generated through the transfer of brain parts to different bodies make no sense, as memories are partly defined by their first-person perspective, and are related to a person's bodily capacities. Due to the different embodiment involved, so-called 'transferred memories' not only do not have a first-person perspective, but cannot do so. This indicates that the claim that one person could remember events in the same way as a different bodied person is incorrect and misleading.

I have also argued that the view that bodily differences are inconsequential to mental life, or to the quality of one's memories, is false and misleading. Each individual body has its own unique structure, capacities, activities, orientation on

the world and so on, such that it makes no sense to assume that the mental life or memories of an individual with one body, could be duplicated by an individual with a different body.

In addition, I have argued that the claim that brain bisections could produce multiple streams of consciousness is unsupported, as this proposition is based on the mistaken claim that brain bisection surgery produces dual streams of consciousness in a single-bodied person, and that the notion of dual consciousness is coherent. Using current neurological findings, I have claimed that while commissurotomies highlight the different contributions made by the right and left hemispheres to a person's stream of thoughts, feelings, and so on, they do not indicate that there is not, in the end, a single, conscious stream, attached to a single-bodied individual. I have argued that because particular streams of consciousness are defined by particular bodies, the results of brain bisection surgery in no way supports the disruption between personal identity and bodily identity. I have also argued that the notion of split streams of consciousness, each with a first-person perspective, emanating from a single body is in itself completely incoherent.

Finally, I have argued that the brain's connection to the body by means of various bodily systems, such as the circulatory system and the nervous system, demonstrates ways in which the capacities and activities of particular brains are intrinsically connected to the capacitates and activities of particular bodies. Consideration of the circulatory system demonstrates ways in which the activities of the body, such as the ingestion of food or alcohol, or the secretion of hormones, are connected to the activities occurring in the brain, and correspondingly, are connected to a person's mental life. Similarly, consideration of the nervous system shows that this system is a vital means of connection between the brain and other parts of the body. Where the activities of the brain

are crucially affected by the nervous system, such as in cases of depression or Alzheimer's disease, mental life is affected accordingly. This demonstrates that what happens in particular bodies affects what happens in particular brains, and correspondingly, in particular persons' mental lives. I have argued that due to the inherent connection between mental life and personal identity, these facts must be acknowledged when accounting for personal identity, if the account is to be in any way accurate or comprehensive. Due to this inherent connection between bodies, brains, minds, and personal identity, it is difficult to see how the disjunction between bodily identity and personal identity, which the psychological continuity criterion maintains, can be supported. In addition to these problems concerning the body, the psychological continuity criterion also has difficulty in accounting for the self. This final topic is addressed in the next chapter.

Chapter 6 The Self as Dynamic Unity

6.1 Introduction

The final area of neglect to be addressed in this thesis concerns the psychological continuity criterion's characterisation of the self. According to the Reductionist conception of the self, selves are nothing more than sets of causally connected mental states, which have no independent role to play in personal identity. Parfit claims that personal identity does not require a foundational or transcendent self, as personal identity is grounded in psychological states alone, and does not depend on underlying entities such as selves or souls. These entities would be superfluous, as they are unessential to psychological continuity, and consequently, unessential also to personal identity. On this view, the self is conceived to be a mere function of psychological continuity, and rather than a crucial determining factor of it. Consequently, when referring to subjects of experience, these subjects are not taken to have any real or tangible existence, other than as linguistic devices or tools of reference, used to refer to aggregations or bundles of thoughts and experiences. One result of this view of is that without a self in which to ground personal identity, progressive and considerable change in a person's psychology can mean that a person can, over time, become a different person. Parfit claims that this means that personal identity is ultimately inconsequential and meaningless (Parfit 1984), pp 219-243. In this chapter, I argue that the Reductionist stance on the self is mistaken, and results from a misconception of the self, its origin and development, its role in experience, and its contribution to a person's *own* sense of identity. In contrast to the Reductionist view, I offer an alternative account of the self, in which the self features not only as an essential component of mental life, but also, as an essential component of the embodied activity on which that mental life depends.

Underpinning Parfit's stance on the self is the assumption that only two kinds of selves are possible. The first possibility is understood as a 'Cartesian soul' or a similar 'unknown entity,' which although unknown, nevertheless grounds experience and mental life. Parfit claims that we have no evidence of this type of self, and that we cannot, therefore, presume that such selves exist. To assume that they do is unsound, and philosophically unacceptable. The second possibility is that selves are formal or linguistic devices, used to describe experience and mental life, but which do not refer to anything that actually exists. This is the view that Parfit accepts. He also holds that without a self in which to ground experience, sufficient change in a person's psychological states could mean that personal identity has also changed. In cases where the amount of change could not be determined, it would thus be an empty question whether a person had one identity or a different identity. It is this conception of personal identity which leads to the view that personal identity itself is ultimately a meaningless concept, and to the view that preference for self-concern over other-concern is irrational. It is unclear, however, how these views can be consistent with Parfit's reductionist stance, since it is hard to see what self-concern or other-concern might mean here.

While Parfit might be right in rejecting the idea of the Cartesian self, his reasoning, however, does not vindicate the second alternative he offers. There are at least two important issues at stake in this regard. The first issue concerns whether the idea of a Cartesian self is indeed coherent or viable. The second concerns whether the two alternatives he offers do indeed exhaust the available possibilities. On the first issue, I largely agree with Parfit, but would argue more strongly that even if a Cartesian self were to exist, it would be of no help in relation to questions concerning personal identity. A self that is involved in personal identity is a self that is involved in the experiences and thoughts from which that identity is constructed. Such a self is knowable, and does not lie outside the

empirical realm. It is not transcendent of experience. The objection to a transcendent self furnished by Parfit, however, is not that it cannot be the ground of experience, but rather, that we have no evidence of such a self. On my view, even were we to have such evidence,[133] it would be of no value here, and would, therefore, be irrelevant. In other words, whether transcendent souls or similar entities exist is not an issue for the personal identity debate. What is at issue is whether an empirical self, a self which is conceived and apprehended in the course of normal experience, is crucially involved in that experience, and consequently, in personal identity.

Regarding the second issue, the Reductionist view, sometimes referred to as the 'no-self' view, is not based on any detailed investigation of what the self might be, how the self might develop, or why the self might be essential to experience and mental life. More specifically, the no-self view has not arisen out of an inquiry into what it means to *be* an experiencing subject, to *know* that one has experience, or to *have* a sense of self or 'self-knowledge.' These issues are crucial to the self, and ought to be investigated if any definitive conclusion about the self is to be reached. As the Reductionist view has not confronted these issues, however, its investigation of the self cannot be considered to have been thorough, or to be a sound basis for conclusive views about the self. This chapter aims to remedy this deficiency, by inquiring into the nature of the self, into the reasons why the self might exist, and into why the self might be essential to experience and mental life. Based on these inquiries, an alternative view of the self to that of Parfit is proposed. According to this alternative view, the self is neither a transcendent entity, nor a mere aggregate or bundle of thoughts and experiences. I argue that the self is a dynamic unity that accompanies all our experiences, that is essential to experience, and that is given in experience. To understand this conception of the self, it is necessary to appreciate that the self at issue here is not a 'thing' in

[133]Such as might be provided by mystics, religious devotees, and similar persons.

the way conceived by many theorists, but is rather a sense of unity within oneself, which is experienced in different ways at different times. This sense of unity is incremental, and begins in a primitive form early in life, and develops through various stages to a more mature and richer sense of unity later in life. In line with this developmental approach to the self, I argue that the self is first experienced bodily in young infants and later cognitively in adults. As this embodied experience of the self is foundational to the later cognitive experience of the self, I also argue that the embodied self is primary to the cognitive self, and that it is only due to the embodied self that the cognitive self becomes possible.

In conceiving of the self as a dynamic unity, it must be made clear here that reference to 'the self,' 'the cognitive self,' 'the embodied self,' 'the dynamic self' or 'the mental self' in the present chapter does not imply that the self is a self-sufficient entity in its own right. While the self is considered to be real and tangible, it is, as will become apparent, a unity of many things, and thus, cannot be considered apart from the things which are contained in that unity. In many instances, reference to the 'self' refers to the 'sense of self,' that is, the sense that one has of oneself as a single, more-or-less united entity, with thoughts and desires, and with the bodily capacity to act on these thoughts and desires. These points will become clearer as the argument for the self proceeds. My case begins by explicating further the 'no-self' view, and then proceeds by outlining the case for the self as a dynamic unity. I then present arguments in support of the need for cognitive unity, and follow this by elucidating the significance of the embodied self. I then draw on empirical evidence to demonstrate the significance of the embodied self to young infants. Finally, I inquire into the cognitive self's development, and explore the relevance to the self of autobiography and narrative.

6.2 The 'No-Self' View

The idea that the self has no evidential basis is not peculiar to Parfit, but is already adumbrated in Hume's work. Hume holds that the self cannot be substantial, as 'substance' itself is a meaningless concept. He claims that we have no 'impression' of substance, and that therefore, we cannot presume that substance exists. Hume similarly claims that no impression of a self is encountered in introspection, and that we therefore cannot claim to have a self. While Hume accepts that most people believe in the existence of a self, he claims that this belief is an illusion. Our ordinary conception of the self is based on a 'fiction.' All that really exists are bundles of fleeting perceptions. These perceptions are related to each other by no more than constancy and coherent change (Hume 1888), 1.4.6. Hume, however, subsequently admits to confusion, as on the one hand, each perception appears to be separate and distinct, and on the other, without a self, there is no principle of unity which binds these perceptions together (Hume 1962), pp 324-331. Thus, although the self is a fiction, unless we maintain this fiction, our experiences could not be coherent. Many theorists since Hume have found themselves equally unable to account for the self, and have therefore rejected the idea of a substantial self, in favour of a 'nominal' self, that is, a self which has no substantial existence outside of particular mental states, or bundles of mental states. Parfit's work provides an example of this conception of the self.

Evidence of the psychological continuity criterion's reductive conception of the self is found in Parfit's thought experiments. One example is the scenario concerning Parfit's imaginary reduplication and transportation to Mars. Developing this thought experiment, Parfit considers the point at which the thoughts of the earthly Parfit are supposedly taken up by the reduplicated entity on Mars. Parfit considers whether an underlying self would be required for this to happen. Relating the possible situation, Parfit suggests that prior to teletransportation, earthly Parfit might think: 'Snow is falling,' Parfit then claims that the Parfitian duplicate would

inherit the memory of this thought at the instance of teletransportation. Parfit then suggests that subsequent to teletransportation, the Parfitian duplicate might follow-up this thought with another thought, such as: 'So it must be cold.' Parfit then claims that because the Parfitian duplicate has 'inherited' the thought: 'Snow is falling,' that the Parfitian duplicate would believe that he, the duplicate, was the person who had this thought in the first place. Parfit claims that duplicated Parfit would be unable to tell from the 'content of his experiences' the difference between whether *he* had the first thought, or whether someone else had the first thought. Parfit draws from this the conclusion that the contents of our experiences tell us nothing about ourselves over and above the experiences themselves. We thus cannot tell from our experiences whether or not we are entities which exist over and above those particular experiences. Any supposed self-awareness which we do have is no more than an awareness of our continuing psychology. Parfit claims that unless we can point to anything in *addition* to this awareness, we have no warrant to claim that anything else exists. And without such warrant, we should accept the Reductionist position that there is no such thing as a self (Parfit 1984), pp 223-224.

The problem with these arguments, however, is that they are merely directed at the idea of the self as an underlying Cartesian entity, and do not give any real consideration to the question of what the self might actually be. These arguments do not address key issues concerning the self, such as other views about the nature of the self, how the self relates to experience, or whether the self is relevant to personal identity. In fact, the above scenario steers us away from the self before the concept itself is examined in any detail. Moreover, the arguments presented depend on the claims they purport to prove. Consider that when the duplicate supposedly inherits earthly Parfit's memories, it is taken for granted that there is no self among the things inherited. Parfit's claim amounts to the proposition that no self goes along with the memories, and therefore, no self is

involved in the memories. I do not see how this argument amounts to an examination of the concept 'self,' or how it produces any conclusive views about whether selves exist or not.

As I explained earlier, the impetus for Parfit's 'no self' view results largely from the rejection of the idea of the Cartesian self, and from the apparent unavailability of any viable alternative view. But as also explained, this apparent unavailability does not justify the complete rejection of the self, and the resort to a bundle theory of experience and mental life. The conclusion that selves do not exist is drawn on the basis of very thin and dubious evidence. Specifically, this evidence is that: 1) there is no evidence of the only viable alternative, namely Cartesian souls, and 2) that selves are not encountered directly in experience. This reasoning, used by Hume, Parfit, and others already assumes a certain conception of the self, and thus does not amount to a serious examination of what the self might be. The problem here is that the apparent failure of viable alternatives to the no-self view, coupled with the apparent inability to directly encounter the self, is taken as *definitive evidence* of the self's non-existence. Scrutiny of this reasoning shows, however, that this conclusion drawn does not necessarily follow from the premises. Another possibility is that the conception of the self which Parfit and others hold needs rethinking. The Reductionist conception of the self follows from certain preconceptions about the self, and from applying such preconceptions to any thoughts about the self, or to any inquiries into the self. But if these preconceptions about the self are mistaken, and if they are taken as the starting point of inquiring into the self, any such inquiry is inevitably doomed to producing faulty and misleading conclusions. A more productive approach would be to rethink the nature of the self, and to discard the 'all or nothing' idea that the self is *either* a Cartesian or similar entity, *or* it is virtually nothing at all. It would be more useful to consider different possible forms of the self, and to examine precisely what the self 'does' for an experiencing subject, which would not be 'done' if selves did not exist. This means we should consider the sense of self from the

perspective of an *experiencing* subject, rather than from that of a merely a *grammatical* one. When the self is considered in this way, it becomes apparent that a certain form of the self is essential to the possibility of experience and mental life, such that without such a self, neither experience nor mental life could occur.

6.3 Searching for the Self

As a topic of philosophical inquiry, the self holds a unique place. Because selves are the subjects of experience, the item being studied is the same item as that which is doing the studying. This means that it may be a mistake to hold strong preconceptions about what the self might be, as this may blind us to recognising the various forms of self-awareness that are available to us at different times. One way that the self might be understood, which is little explored by Parfit, is to consider the self as the unity of the bodily and mental capacities of a single individual, together with, and modified by, that individual's first-person knowledge of these capacities. On my view, this integrated unity is essential for either thought or experience to occur, as, unless there was some form of meaningful connection between these various facets of our experiences and capacities, mental life would not be coherent, and action could not be initiated. This view of the self may initially seem puzzling, and somewhat at odds with traditional views of the self. But when this view is considered in more detail, it becomes evident that the sense we have of ourselves as individual entities is very much tied to the notion of ourselves as entities with particular bodies and bodily capacities, and with particular sets of thoughts. This idea is supported by the fact that where persons' capacities and knowledge are well integrated, these persons have a strong or robust sense of self.

Conversely, when persons' capacities or mental life are adversely affected, such as through bodily injury or mental trauma, the sense of self becomes depleted and

shrinks. In severe cases, such as grave physical damage or illness, or complete mental breakdown, the sense of self may almost disappear altogether. Such persons lose their sense of identity and are unable to know how to act, or what they should do next.

Considering the self as a dynamic unity has at least four important aspects. The first has already been referred to, but bears repeating. It is the point that the self is not a 'thing', but is a unity of knowledge and capacities we have of ourselves at any given time, which underpins our thoughts and actions, and which is necessary for us to experience the world around us in any form. The second aspect is that the self is developmental; it begins primitively early in life, and develops incrementally throughout life. Third, is that the self provides different forms of self-knowledge from a first-person perspective. And fourth, is that the self is first experienced bodily in young infants, and develops its cognitive aspects subsequent to, and conditional on this earlier bodily experience. Neisser's models of self-knowledge are useful to understanding this developmental approach to the self.

Neisser examines the self in terms of five different types of self-knowledge, each of which represents a stage in the self's development. The most basic self is the *ecological* self. This is the self as it relates to the immediate environment. It is directly perceived right from birth, and is circumscribed by the body's capacity for activity. The ecological self does not always coincide with the body, although it could do so. It might also include things within the body's control field, such as clothes, prosthetic limbs, or even a person's car (Neisser 1988), pp 37-41.

Second, is the *interpersonal* self, which is also directly perceived. This self comprises unreflective, intersubjective relations between at least two persons. Interpersonal selves engage through words, expressions, signals and cues, such

as occur between mothers and young infants. In normal cases, the interpersonal self is present at birth (Neisser 1988), pp 41-46.

Third, is the *extended* self, which is the self that engages in self-reflection. It thus extends beyond merely the present context. It includes the self of the past and the expected self of the future. It is known mostly through autobiographical memory (Neisser 1988), pp 46-50.

Fourth, is the *private* self, which is a self independent of an individual's present circumstances. This self is concerned with inner perception, dreams, intentions, memories, imagination, and other forms of introspection (Neisser 1988), pp 50-52.

Finally, the *conceptual* self is the view we have of ourselves, based on facts we know about ourselves, such as our profession, family relations, cultural standing, nationality, and even how others see us. It concerns those aspects of ourselves that we see as most important and identity-determining. Neisser sees the conceptual self as essentially comprised of at least four of the preceding aspects of self. It incorporates our various forms of self-knowledge into the view which we have of ourselves, and which we assume others to have of us. Our self-concept is important, as it is the overarching concept through which we know ourselves, and through which our diverse aspects are united into a single person (Neisser 1988), pp 50-54.

Neisser conceives the self as a developmental phenomenon, manifest first in bodily form, and later in a cognitive form. The ecological and interpersonal selves are realised through embodiment, and the later extended, private, and conceptual selves are realised through cognitive development. Similarly, the dynamic conception of the self is construed in terms of embodied unity and cognitive unity. These are not completely discrete unities, but are progressive stages in the

development of a numerically single compound unity of knowledge and ability.[134] To better understand this developmental approach, the self is now examined in more detail. Because the cognitive aspects of the self are those which most concern the personal identity debate, this aspect of the self is addressed first. Consideration of the cognitive self is followed by a detailed investigation into the self's earlier manifest bodily aspects. The bodily self is considered in some detail, as I believe that understanding the bodily self is crucial to appreciating the self's role mental life and personal identity, and also, because the connection between the bodily self and mental life is little understood in the personal identity debate, particularly as it relates to the psychological continuity criterion.

6.4 Cognitive Unity

Cognitive unity is the unity of accumulated thoughts and experiences that is required for such thoughts and experiences to make sense. Cognitive unity is a form of self-unity, as the self just *is* the unity of thoughts and experiences, and of the capacities to act on these experiences. To distinguish the cognitive aspects of the self from other aspects, the self is here referred to as the *cognitive* self.[135] One way in which the cognitive self is evident is in the first-person character of experience. For, unless I have some knowledge of my thoughts and capacities, I cannot intentionally act on them.[136] The importance of the self to mental life has been addressed by Kant and more recently by Malpas. Their work was considered earlier in reference to mental holism.[137] The role of the self in making

---

[134]'Knowledge' here is used in its most general sense, and refers not only to conceptualised knowledge, but also to unreflective bodily knowledge, such as the innate, unselfconscious knowledge of where my mouth is.

[135]When referring to the 'cognitive self', 'experience' will be understood as conceptualised, self-aware experience, such as contributes to, and is informed by mental life. When referring to the 'embodied self,' 'experience' will refer to non self-conscious embodied experience, such as quenching thirst, moving limbs, or crying from hunger.

[136]This awareness must be understood in a general sense, and includes awareness of our possibilities. I may be unsure whether I can lift a heavy object until I try, but it must be a real possibility for me to intend and attempt the action.

[137]The work of Kant and Malpas was addressed in Chapter 4.4. While this work is referred to again here, the previous emphasis was on the implications of the predominantly holistic character of mental

possible experience, and the holistic character of mental states are integrally related concepts, and need to be understood in light of each other.

Recall that Kant was concerned to set out the necessary conditions for the possibility of experience. More generally, we might say he was concerned with establishing the conditions for the very possibility of contentful mental states. Such states depend upon the synthesising activity of the understanding, since mental states, and the content of those states, is essentially dependent on the connections ('synthesis') between those states and their elements. Moreover, such connections are themselves worked out and expressed, as we saw earlier, through the way in which all such states are referred back to a single, unitary consciousness - a single 'I.' Without the 'I think', claims Kant, experience would be 'nothing to me' (Kant 1929), B 132.

Apprehending experience from a first-person perspective means that with respect to any and every experience, that experience must be apprehended by some experiencing subject, and thereby connected up with other such experiences, mental states and so forth. This is to say simply that the having of experiences, or the having of mental states, depends upon there being a system or network of such states, and for there to be such a system is for there to be a single subject to whom those states 'belong.' The idea of the single subject is just the idea of there being states that are appropriately connected or 'unified.' Unless thoughts were present to a single consciousness in this sense (and note that this does not imply the existence of a consciousness or 'self' that is independent of and separable from the states that are present to it), it is difficult to conceive what a thought could be, or how such a thought could be understood. For Kant, a single consciousness is a *unified* consciousness, that is, one in which thoughts and other mental items cohere into a more-or-less integrated unity, since, unless thoughts

---

states, whereas here, the emphasis is on the unitary structure of the self, particularly as it relates to mental coherence, and to the capacity for action.

were unified in this way, they would be independent, isolated, solitary, and consequently, meaningless to the person whose thoughts they were - they would, in fact, be contentless and thus not be mental states at all.

Since he takes this unified subjectivity as a necessary condition of experience, Kant concludes that the mental unity at issue here is *transcendental* to, or is an *a priori* condition of experience.[138] Designated the 'Transcendental Unity of Apperception,' this unity is the unity of thought required for the comprehension of experience, and for me to know that my experiences are mine (Kant 1929), B 132. Kant claims that although this transcendental unity is an *apriori* condition of experience, we can know nothing of it beyond its formal necessity, and thus cannot attribute characteristics to it, such as those of an immaterial soul or similar unknown entity (Kant 1929), First Paralogism, A 348-351. It is, as I noted immediately above, not a unity that can be taken to imply some substantial and independent self or subject, but rather refers us to what might be viewed as an almost purely 'formal' unity - a unity that is identical with the connectedness of the states themselves, and that is worked out through the interconnection between states. This unity is established in concrete experience, and, keeping in mind the arguments from Malpas that we have already considered, also in embodied action.

Kant's view that self-unity is a pre-requisite to the apprehension of experience has important implications for the account of the self. It draws attention to the point that experience does not occur as a matter of course, but requires the presence of an appropriately constituted unifying structure. For Kant, this unifying structure depends on an *apriori* unifying principle, exemplified in the *Transcendental Unity of Apperception*. The fact that this unifying principle is a necessary condition of

[138]'Transcendental' as referred to here must not be confused with 'transcendent' referred to earlier. I take 'transcendental' to refer to a necessary precondition of something, that is, in the Kantian sense, and 'transcendent' to refer to something which exists outside of or beyond the material universe.

experience, such that without it, experience would not be possible, must not be confused with the similar, but quite different claim, that experience depends on entities existing *transcendent* to, or outside of, experience. As has been stated, such transcendent entities, were they to exist, would not provide the necessary unifying principle *within* experience. So, although Kant points to the need for apperceptive unity, that is, the unity required for *me* to know that my experience is *mine,* this unity is not to be thought of as existing apart from the occurrence of experience itself. Apperceptive unity can best be understood as 'hovering,' as it were, at the outer limits of experience, being both a necessary precondition of experience, but as arising only out of experience. The work of Malpas makes this position clearer.

In his account of the holistic character of the mental, Malpas argues that mental unity is tied to the apprehension of experience by a single subjectivity. His emphasis on the relation between the first-person apprehension of experience and the grasp of oneself as an experiencing subject was referred to earlier. Malpas argues that the existence of a unified subjectivity or self prior to the occurrence of experience makes no sense, and merely defers the question of unity. He claims that if the subject which experiences is a subject which exists *outside* the experiential realm, such a subject would be incapable of conferring unity on experience. Unity, after all, is not a mere causal affair, such as might be found in the operation of brain mechanisms (or in Parfit's thought experiments). These connections are no more than physical connections, and do nothing to explain the coherence of content or meaning (Malpas 1999), p 88. For example, knowing that certain principles of hygiene contribute to the prevention of disease involves more than having certain neurones firing in my brain and making certain synaptic connections (although having the thought that: 'hygiene is important to preventing disease' might be identical and coincident with the firing of certain neurones in my brain). This knowing also involves my understanding the hygiene

principles involved, how these principles relate to the prevention of disease, and how the absence of these principles could cause disease to occur. The unity of thought is thus a unity of content, and therefore, can only be derived from the realm to which that content refers. Contentful connections between beliefs and so on, to be intrinsic to those states, cannot simply be conferred by some entity to which those states are attributed. The unity at issue is a 'rational' unity.

For Malpas, the unity of the self is a matter of the contentful unity of mental states and is thus a unity created *by* experience, rather than being a unity which exists prior to the occurrence of experience. He sees that such a unity is a direct consequence of the holistic character of mental states, and, as being integral to the identity of those states, and consequently, also integral to the identity of the self. Thus, part of what makes *my* self the self that it is, is the fact that I have certain mental states, which, being present to me in a single consciousness, I can identify as mine. Similarly, part of what makes those mental states the states that they are is the fact that they are present to my consciousness, and are interconnected with other states which also belong to me (Malpas 1999), p 89.

While the self at issue here is an essential contributor to cognitive unity, it is not merely a 'mental' self. Malpas argues that underpinning the capacity for cognitive unity is the capacity for intentional action. This capacity involves an understanding of oneself in relation to other objects in the world, the ability to affect those objects, and an appreciation of that ability.[139] Implicit in understanding one's relation to other things is a grasp of spatiality, that is, the recognition that one is a three-dimensional object in a world of other three-dimensional objects, all of which are located in particular places, and which hold particular spatial relations to each other. The occurrence of action requires that these various elements are related

[139]For example, having the ability to act, and having an awareness of that ability is important, as unless I could move my fingers in a certain way, and to know that I could so move them, I would be unable to lift a cup to my lips to take a drink.

to each other in the appropriate way. More specifically, there is an integral connection between the self, spatiality, and agency, or, as Malpas puts it, between the subjective and objective elements of experience:

> So in establishing the necessary connection of self and content with spatiality and agency, a connection is also indicated between these concepts and the concept of objective space, and more generally, with the idea of objectivity. What is thus starting to appear is a structure in which subjective and objective elements are interconnected and interdependent. It is on the basis of this structure that thought and experience are possible (Malpas 1999), pp 99-100.

The unity of the self is thus a unity that obtains between the subjective and objective elements of experience. This means the self can be neither a transcendent entity, nor a mere disunited bundle. A transcendent entity does not concern experience, and a disunited bundle lacks any principles of unity or integration. Neither of these selves are viable options, as the second cannot explain how experience has coherence and meaning, and the first cannot explain how experience occurs at all.

Malpas points out that experience requires certain essential components, and that these components must stand in relation to each other in the appropriate way. It is not just a matter of having, for example, the required neural structure of the brain, or a world of external objects (although these are not irrelevant), but is a crucial matter of having efficacious connections between these components, and of having certain bodily capacities for action. It is also a matter of perceiving these components of experience from a first-person perspective, or, put another way, of having apperception. The self could thus be understood as the apperceptive unity of experience, which is not transcendent to, but is immanent in experience, and which is required for an agent to have experience, and to be able to act on that experience.

Consider, for example, whether I could initiate action or have coherent thoughts and experiences if there were no unity between the elements which composed my thoughts and experiences, or if I did not see myself as standing in certain relations to these elements. I would not, for example, be able to eat unless there was the appropriate coherence between the feeling of hunger, the desire for food, the knowledge that food would satisfy my hunger, the knowledge of where and how to obtain food, the capacity to move my limbs in the appropriate manner to obtain food and to eat it, the first-person knowledge of these capacities, and, most importantly, the sense that it was *me* who was the subject of these thoughts and activities. Were all these elements to exist individually, disunited and isolated from each other, they would not be the elements that they are, and the activity which would result from their unity would not be possible.

The unity which makes thought and experience possible is the unity of the self, as self-unity just *is* the unity of the elements required for thought and experience to occur, and to be present to a single subjectivity. This means that the self is essential to all experience, and can, therefore, be taken for granted to accompany all experience. Put another way, because the unity of the self is part of what it *means* to have experience, the very fact that experience occurs at all means that the self is already present. Experience without the self or self-unity does not make any sense. Indeed, once we see things this way, then we do see that the self is not some entity over and above experience and actions, it is the unity of experience and actions that makes experience and action possible.

The unity of the self is not just a unity of mental elements, but, due to its involvement in action, also includes the bodily capacities and abilities of the agent concerned. There is thus a close connection between a person's sense of self and a person's mental and physical abilities, and the degree to which these abilities are developed. Self-experience is thus not static, but is open to change and

development, in accordance with a person's particular activities and engagements with the world, and with objects in the world. This means that although, as Neisser suggests, the self has a private, first-person aspect, the self also has externalised aspects, which, through engagement with the world of objects, contributes to the self's modification and development.

The psychological continuity criterion does not recognise these aspects of the self. According to this criterion, the self is a kind of non-entity, with no significant causal power to affect or contribute to mental life, or to the experience from which mental life is derived. This means that the capacity for thought and experience is not acknowledged as requiring preconditions, and that the possibility of thought and experience, is, therefore, taken for granted. But as shown above, this conception of thought and experience is inaccurate. Experience and mental life require coherence, and coherence requires unity within a single subjectivity. The unity of experience and mental life within a single subjectivity is the unity of the self, as it is only this unity which includes all the elements of which experience and mental life are composed. Thus, in leaving the self out of personal identity, the psychological continuity criterion neglects to account for the unity which is essential to experience and mental life, and in doing so, leaves crucial aspects of personal identity unexplained. To better understand the foundations of cognitive unity, the bodily unity which underpins mental life is now explored.

6.5 Embodied Unity

Embodied unity refers to the sense of one's body as a more-or-less united entity which under one's control. The experience of one's body as a unity is an important aspect of self-experience, and is thus a crucial aspect the *embodied* self. The embodied self broadly refers to the non-conceptualised self-experience which is encountered through the body, directly and unmediated. This is a more 'primitive' aspect of the self, as it is unreflected and unself-conscious.

Experiencing the self through the body is an important condition of action, as unless one has a sense of oneself as a more-or-less unified body, one does not have the control of one's body or body parts required to initiate or perform action. Because performing action is essential to mental life, the embodied self is a necessary condition of the cognitive self. For it is only in virtue of the bodily activity which the embodied self makes possible, that the mental life of the cognitive self can occur. The embodied self is thus a precursor to the cognitive self, and is the first form of self-experience.

The integral link between conceptualised experience and embodied activity was referred to earlier. Only creatures with the capacity for action are creatures who are capable of thought and experience. As Malpas has pointed out, activity is integral to the unity and coherence of mental life. A crucial feature of activity is the ability to orient and direct one's body in space. This requires the ability to distinguish between oneself and the rest of the world. Malpas notes that this distinction marks the difference between the subjective and the objective realms of experience, sometimes understood as what a creature can do (and knows it can do), and what it actually does.

Important to knowing the difference between oneself and other objects, and of appreciating one's capacities, is the 'body-schema.' According to Malpas and others,[140] the body-schema is the innate, subconscious knowledge which we have of our bodies. This includes such things as knowing that our bodies are separate from other objects, where our limbs are, and the sense of our orientation in space. I seem to know without thinking, for example, whether or not my legs are crossed, or where my hands are, or whether I am sitting or standing. The body-schema

[140]For example, see John Campbell and Shaun Gallagher. See *'The Body Image and Self-Consciousness'* and *'Body-Schema and Intentionality'* in (Bermúdez, Marcel et al. 1995), pp 29-42; pp 225-244.

permits us to be aware of, and to monitor changes in our bodies and in the orientation of our bodies Malpas 1999), pp 109-111.

The body-schema is an important aspect of the embodied self, as it provides first-person knowledge of myself as an object in space and enables me to differentiate myself as a real, tangible existent from the objects around me. The body-schema is not synonymous with the embodied self, but is an essential component of it. For unless I had a sense of myself as an object separate from other objects, and of having limbs, and of being oriented in space, it is difficult to perceive how I could recognise myself *as* myself, or as an entity with the potential for movement, or with the ability to effect change in the world.

Also involved in the embodied self is the body-image. This refers to the reflective, self-conscious awareness or beliefs about one's body, and is the sense of one's body which is involved in the actions which one may commit at a particular time, rather than those of whose potential we are more generally aware. As the potential and importance of the body-schema and the body-image are relatively new in the study of the self, the difference between their roles is only incompletely understood. But what is apparent from studies of the relation between embodiment and the self,[141] is that a direct relation pertains between the sense of one's bodily capacities and one's orientation in space, and the sense of one's self as an efficacious object, capable of effecting changes in objects separate from oneself.

Not all theorists, however, recognise the primacy of the embodied self. Galen Strawson, for example, argues that the self is primarily mental :

> the self ... is ... a mental presence; a mental someone; a single mental thing that is a conscious subject of experience, that has a certain character or personality, and that is

[141] Such as the studies mentioned earlier (see previous footnote), and those which will be referred to in the remainder of this chapter.

> in some sense distinct from all its particular experiences, thoughts, and so on, and indeed from all other things. It is crucial that it is thought of as a distinctively mental phenomenon (Strawson 1997), p 407.

Strawson argues that bodies are inconsequential to the self, as they are little more than 'vessels' which contain the most essential *mental* part of ourselves (Strawson 1997), pp 405-414.

George Butterworth is one theorist who takes issue with Strawson, arguing that the mental or conceptual self is underpinned by a more basic self, which is experienced through the body, and which first arises as a perceptual response to the immediate environment:

> Perceiving is a spatio-temporal process which provides a continuous flow of information about the embodied self in its encounters with the physical and social world (Butterworth 1998), p 132.

Butterworth claims that the embodied self is the self's most fundamental aspect, and is foundational and intrinsic to the more developed, cognitive aspects of the self. Referring to Strawson's mental self, Butterworth states that an 'overly cognitive' view of the self divides the self from the environment and from the body. But, argues Butterworth, this is a mistake, as it is only in virtue of the embodied self that the cognitive self can develop. The distinction which Butterworth makes is between the immediate perceptual experience of the self, and the formation of a self-concept. In opposition to Strawson, Butterworth argues that immediate awareness of the self occurs before the more reflective awareness of self-concept formation (Butterworth 1998), p 132-133.

Butterworth refers to Neisser's five kinds of self-knowledge, in which the ecological self, that is, the self as experienced through the body's interaction with the immediate environment, is understood to be the primary form of self-experience, and which, with the interpersonal self, is the natural precursor to

concept-formation. For it is only in virtue of the body's interaction with its immediate environment that the information on which concepts are based is received (Butterworth 1998), p 133.

Butterworth's view highlights the inadequacy of a purely mental self by pointing out the significance of the embodied self to the possibility of the later conceptual self, and to the consequent maturation of mental life. For unless we have a sense of ourselves as bodies located in space, co-existing with other objects, receiving perceptual information from those objects, it is difficult to see how our thoughts and concepts about those objects could be formed, or how such thoughts and concepts could have any coherent meaning. It can thus be seen that the embodied sense of self is a crucial aspect of mental life, which needs to be recognised by the psychological continuity criterion, if its account of personal identity is to accurately reflect the essential features of mental life and activity of which personal identity is composed.

As well as holding metaphysical primacy to the cognitive self, the embodied self also holds temporal primacy. For the embodied self is the first way in which young infants experience the self. This is apparent from the study of infant behaviour, in which the co-ordinated bodily activities of infants are observed. These studies are comparatively new in the study of the self, and thus add new and important information to the traditional philosophical considerations of the self. Important among the findings of these studies is the view that the sense of self as a unity is present in infants from the earliest experiences of life, and that as an infant's capacities and experiences progress and develop, the infant's sense of self-unity progresses and develops also. This view of the self is in contrast to Parfit's view, which sees the self as a peripheral concept to experience, and as, therefore, irrelevant to personal identity. The view of the self presented here, however, takes the self seriously, and conceives it to be essentially the first-person perspective of

self-unity which is necessary for experience of any kind to occur, whether such experiences are those of young infants, or whether they are those of fully-functioning cognitively developed adults. As will become more evident as the arguments for the dynamic sense of self proceed, because this sense of self is integral to all forms of experience and mental life, it is inevitably also integral to personal identity. To develop the arguments for the embodied self and to show how this self is primary and fundamentally necessary to the cognitive self, the embodied self is now considered in more detail, first by considering the relevance of bodily movement, and second, by considering the relevance of perceptual integration.

6.6 The Importance of Movement

Recent studies of young infants support the view that the embodied self is primary to, and preconditional to the later developing cognitive self. These studies indicate that the self is experienced right from birth, and that it develops throughout infancy, along with the maturation of the brain's neural structures. Because they experience this bodily sense of self, infants are able to differentiate themselves from the rest of the world, to learn about their own bodily capacities, and consequently, to develop the skill required to orient and direct their movements.

Maxine Sheets-Johnstone supports the view that infants first experience their sense of self through movement.[142] She argues that through their bodily movements, infants learn things about themselves, such as that they are not rooted to the ground, that they have limbs, and that the world is made up of ground and space. Activities such as bending, stretching, and lifting give infants a

[142]An objector could claim that quadriplegics have a sense of self, yet cannot move in the normal way. In most cases however, these persons would have developed in the normal way prior to their disability. I know of no cases in which persons have been totally paralysed from birth. Were such a thing to be possible (which I doubt), such persons might develop a minimal sense of self through their ability to blink, breath, move their tongue, swallow, and so on. Were none of these movements possible, it is difficult to see how such an individual would qualify as a human being in the normally understood way.

sense of themselves and their surroundings, and of things such as closeness, distance and direction. Sheets-Johnstone argues that movement is the most fundamental source of knowledge:

> Coming to know the world in a quite literal sense means coming to grips with it - exploring it, searching it, discovering it in and through movement. There is no human culture in which movement is not epistemologically central in this way. There is, indeed, no culture in which movement is not our mother tongue (Sheets-Johnstone 1999), p 226.

Movement is thus likely to be the earliest source of self-knowledge. Sheets-Johnstone argues that for infants, movement is foundational to other types of learning about themselves, and about objects outside themselves.

Adults are also sensitive to movement, as we often respond to things such as facial expressions, eye and lip movements, and other gestures. Sheets-Johnstone claims we are 'kinetically attuned to each other,' and that the capacity for such attunement is probably innate. Movement teaches both infants and adults about the difference between objects and the perception of objects. Movement allows us to directly engage with the world, rather than to be passive observers (Sheets-Johnstone 1999), pp 228-229.

In these comments, Sheets-Johnstone points to something very important about ourselves, something which we take for granted. This is the fact that there is a fundamental difference between the preconditions of *directed* movement, and the preconditions of the passive reception of movement. Consider, for example, the difference between, on the one hand, being carried from one room to another and someone taking hold of your hand and guiding it towards an object, and, on the other hand, you walking into the room and you moving your hand yourself. In the first instance, you did not initiate or direct the movements, whereas in the second instance, you did. In other words, in the first example, you were passive, and in

the second, you were active. To be passive requires no effort on your part, but to be active does. The effort required for activity is dependent on certain elements, and on there being a certain unity among those elements. The elements concerned include your sense that your body is *your* body, that your body is a united entity (that is, that the parts within it are connected to and integrated with each other), and that your body is under your control. It is the unity between these elements which compose and provide the sense of yourself as an embodied entity and which give you the capacity for directed action.

It is important to realise, however, that the bodily sense of self referred to here is not a self-conscious awareness, or in any way a cognitive awareness. It is much more direct and unmediated than that. Cognitive awareness, in fact, would be of no use to the capacity for embodied activity in the sense being discussed here. Cognitive awareness of the ability to act is secondary to embodied awareness, and develops *out of* embodied awareness. As infants develop, their motility increases, along with the development of the appropriate neural structures in the brain. As conceptual frameworks develop, so also does the ability for self-awareness, and ultimately, for cognitive awareness of their innate, already developed physical abilities.

The innate bodily sense of self, experienced by young infants, is evident from observation of their movements. Infants have a natural ability to move their hands and kick their legs. Studies of these activities show that the infants' movements are not aimless and random, but display a considerable degree of purposeful activity. Video-taped studies of young infants bear this out. A common early practice of young infants is to attempt to put their fists in their mouths. This activity, far from being uncoordinated, involves the correlation of limbs and body parts, all at the same time as much writhing and kicking is going on. In one study of these activities, many infants made contact with the mouth easily, but some

struggled at first. But these infants persisted, gradually improving their aim, until they made the desired contact. Such persistence do not sound like the activities of a random, aimless, disunited, amorphous blob of flesh and bone, but rather like those of an intensely determined and directed being. Studies like these lead some theorists to believe that much self-directed activity occurs before birth, while the infant is still in the womb. George Butterworth, for example, claims that studies like the one above demonstrate that infants have a 'well organized system, which may have benefited from practice in utero' (Butterworth 1995), p 91.

As infants develop and experience more of the world, their sense of self develops also. Information received by infants from their various sensory modalities contributes to this development. Visual perception is one prime example of this process. When combined with movement, visual perception provides *self-specifying* information for normally-sighted infants.[143] This refers to information which permits the infant to distinguish between things which are part of himself, and things which are not. As vision develops, infants learn about objects in their visual field, and about how their own movements make these objects appear to change location in relation to their own position. One way of understanding this process is to consider that the act of perception involves receiving two types of perceptual information, that which relates to the perceiver, and that which relates to the object being perceived. Drawing on the work of Sherrington and Gibson, Butterworth refers to the first of these types as *propriospecific* information and the second as *exterospecific* information. Propriospecific information specifies the subject to himself, and is thus 'self-sensitive.' Exterospecific information specifies objects in the subject's environment. These two types of information are not received separately, but are two simultaneous aspects of a single perceptual response (Butterworth 1995), p 89.

---

[143]Congenitally non-sighted infants do not have the advantage of visually-provided information. Deprived of this aid to the development of the self, their sense of self matures more slowly than that of sighted infants.

Self-specifying information is received by infants as they perform various activities, such as crawling, walking, and running. During this process, infants become increasingly aware of themselves as single, coordinated, individuals. Their movements contribute to their awareness of the boundaries between themselves and other things, and hence of the *difference* between themselves and other things. Infants' instinctive awareness of themselves against their surroundings has been demonstrated in room-moving experiments. The infants' responses to these experiments indicate an innate sense of themselves and of their own bodily unity. In one such experiment, a three-sided 'room' was constructed and suspended above the ground. Infants, sitting down, were placed inside, and the room was then moved back and forth, making it appear to the infants that the end wall was moving backwards and forwards in front of them. Because walls do not normally move in this way, it appeared to the infants that it was they who were swaying backwards and forwards, rather than the wall in front of them. This incongruity made the infants try to correct their posture (which was not, in fact, faulty), and they, therefore, fell over. (Butterworth and Harris 1994), p 82. Butterworth sees these attempts at posture correction as indicating that the infants had an implicit, unself-conscious bodily self-awareness. For, argues Butterworth, it is only if an infant senses that it is he *himself* who is falling over that he would be able take the compensatory action required to try to prevent this from happening (Butterworth 1998), pp 133-134.

The above studies are examples of ways in which infants respond bodily to their surroundings. More specifically, the infants' attempts to orient their body parts *in response* to their surroundings demonstrates that the infants innately sense their separateness from their surroundings, and that they have the capacity to act in response to that separateness. The capacity to respond to one's surroundings with movement coincides with Neisser's concept of the ecological self. It is this early aspect of the self which first receives information about the world, and about

one's place within that world. This information is foundational, as it is only by having the sense of oneself as an embodied creature, capable of moving in one's surroundings, that other more advanced forms of information can make any sense. For, without the sense of myself as a body in space, capable of moving in space, how could I understand what it means for objects other than myself to exist, or what it means for me to perceive them? Without this sense of self, we could make no sense of experience, or of objects, or indeed, of anything at all. This view of the embodied sense of self contrasts with the view that the self has no force or efficaciousness. It points to the fact that the ability to have experience is not a given, but requires certain preconditions of self-awareness, even though for young infants this self-awareness is unconscious, it is nevertheless present in the achievement (and often struggle for) the self-unity which is necessary for directed activities and movements to occur. As this embodied self is an essential precursor to, and an important part of the more developed cognitive self, it is also an essential precursor to mental life, and thus should be recognised when accounting mental life and for personal identity. The primary role of the body in self-awareness is further elucidated by considering the propensity to integrate various facets of perceptual information.

6.7 The Importance of Integration

In addition to motility, the ability of infants to integrate various facets of their experiences is taken as evidence of a bodily sense of self, and in many instances, as evidence that this sense of self is innate. Recall that the self is understood primarily as a unity, in which various aspects of experience and ability are co-ordinated into a single, coherent whole, which is apprehended and perceived by the subject as belonging to and referring to *him.* The ability of infants to integrate experiences includes the ability to connect past experiences with present ones, and the ability to connect various facets of a single experience, so that experience is apprehended as a single, coherent event. This integration process is possible

due to the fact that all our experiences are self-referencial. This means that I do not experience my movements, my perceptual intake, or my thoughts naively or neutrally, as if they were in a vacuum, and did not apply to any particular person, or any particular thing. I experience them very much as *mine,* as belonging to *me.* This sense that what I experience is mine, is only possible due to the unification of the different aspects of a perceptual event into a coherent whole, or, in other words, due to the unity of the self. Experiencing my perceptual responses as mine, and experiencing them as coherent, are thus interrelated and mutually supportive - while each is dependent on the other, neither is primary to the other. The natural inclination to integrate perceptual input was demonstrated above when referring to room-moving experiments. When confronted with perceptually incongruent situations, compensatory action is taken so that a coherent response is reached, which in this case was falling over. When older infants or adults are involved in room-moving experiments, their compensatory response consists of altering their body posture in order to avoid falling over. Many different types of perceptually incongruent situations are possible. In cases where these situations are difficult to resolve, we have the sense that something is wrong, which needs to remedied if perception is to be coherent. The drawings of Max Escher is an example of this type of phenomena.

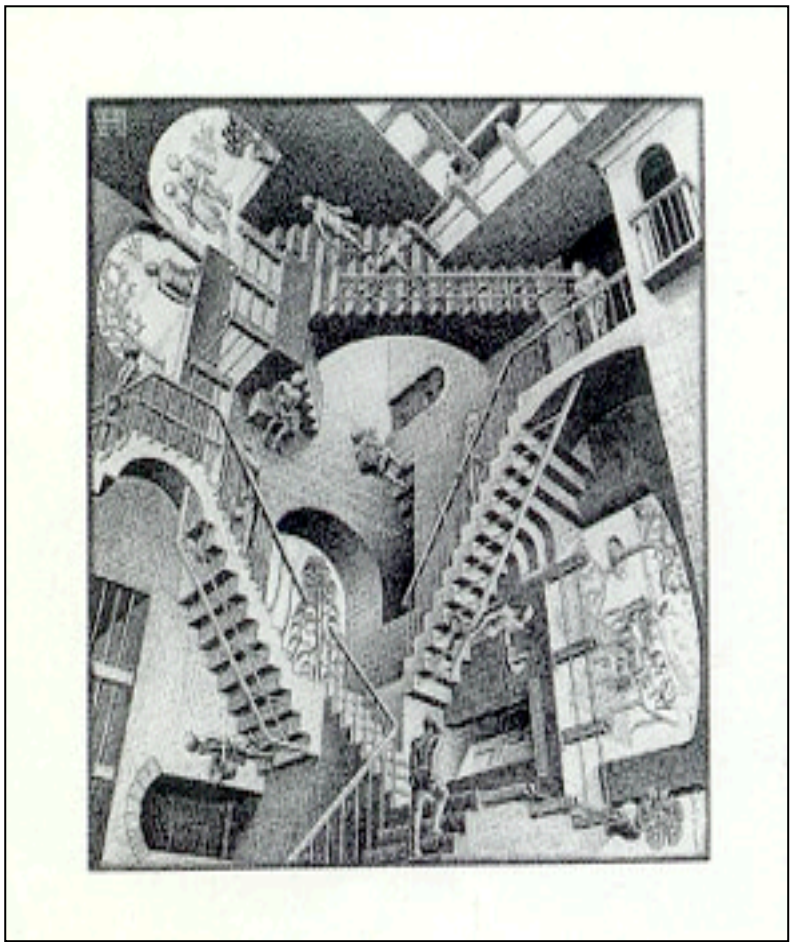

Figure 13 Relativity
Reproduced from http://www.artistsuk.net/acatalog/ARTISTS_UK__PRINTS_133.html

In these drawings, situations are represented which are impossible in the real world, such as perspectives which are distorted and impossible (See figure 13), three-dimensional objects emerging from two-dimensional surfaces, staircases which appear to ascend, but which reach a level seemingly lower than the one from which they started, or objects with incongruous structural features, such that apparently external boundaries also appear to be present inside objects. The reason we find these drawings so confounding is because we instinctively try to integrate the conflicting elements into a coherent whole, and find we are unable to do so. Our instinctive drive to integrate seemingly incompatible elements into a single perceptual response stems from the from the need to unify diverse elements of our experiences into a single subjectivity, so that these elements make sense to that subjectivity. This means that one way of locating the sense of

self is by ascertaining whether the elements of an individual's experience are more-or-less integrated into a single, coherent whole. When considering young infants, this strategy is particularly useful, as it is a reliable indicator that a sense of self is present. Researchers have recently examined the perceptual and bodily responses of young infants in a variety of ways. Their findings indicate that a bodily sense of self is unequivocally present in young infants, and, according to some researchers, is innately present. The innateness of this sense of self contradicts the work of some earlier researchers, such as Maurice Merleau-Ponty.

Merleau-Ponty agrees that the primary mode of self-experience is through the body (Merleau-Ponty 1964), pp 13-125. He notes that before they are able to engage in language, infants exhibit bodily responses, such as smiles and gestures, to those around them. He sees these forms of communication as ways in which we become aware of ourselves, at the same time as we become aware of others. According to Merleau-Ponty, this awareness first begins in infants at three to six months of age. Prior to this time, infants live in a primitive and undifferentiated world, in which there is no self-awareness. Not until about six months of age, claims Merleau-Ponty, can an infant 'look another child in the face,' and know that he is perceiving another person. Merleau-Ponty holds that it is at this age that the myelinzation of the required neural connections is sufficiently advanced for these connections to operate, and for the child to have a sense of its own body schema (Merleau-Ponty 1964), pp 13-125.[144]

More recent research, however, and the devising of new study techniques, causes several theorists to disagree with Merleau-Ponty, and to claim that a sense of self is experienced much earlier that Merleau-Ponty thought. This claim is based on studies of very young infants, including newborns. An important aspect of these studies is the testing of infants' ability to co-ordinate various

[144]'Myelinization' is the developmental process by which neurones in the peripheral nervous system become covered with a layer of fatty membrane (Curtis 1983), p 1101.

perceptual elements into a single, coherent perceptual response, and thus to determine the presence of rudimentary self-awareness. Habituation, olfactory discrimination, phantom limbs, and imitation are some of the topics studied.[145] These studies bear out that the sense of self is present very early in life, and is integral to even the most primitive experiences. This adds further weight to the argument for the necessity of the self to all forms of experience, and to the priority of the embodied self over the more mature cognitive self.

6.7.1 Habituation

'Habituation' is the familiarisation process in which infants are presented with objects which elicit a perceptual response, visual response being an example frequently tested. In such tests, infants are found to initially respond to new objects with great interest and to spend considerable time looking at them. As the objects become familiar, the infants gradually lose interest and look at the objects less. Then, when new objects are presented, the infants again look longer the new stimuli. This variation in response is taken to indicate the infants' ability to discriminate between familiar and unfamiliar objects. When presented with a smiling face, for example, an infant will first look with steady interest, but will eventually lose interest and look away. But when the smiling face is replaced by a 'surprised' face, the infant resumes interest and continues to look at the new stimuli. This is taken to indicate that the infant can tell the difference between the two presentations (Stern 1985), pp 39-41.

The ability to discriminate between earlier and later presentations is only possible if the infant has the capacity to connect (albeit unconsciously) its various

[145]The ability for 'cross-modal translation,' that is, to translate one mode of perception to another, such that one instinctively knows whether what one sees is the same as what one feels, is another infant response which has been tested. See (Gallagher and Meltzoff 1996), pp 223-227, and (Stern 1985), pp 47-51. This ability (for sighted infants) is seen as indicating the presence of an innate body schema, and of the ability to connect different modes of perception in a single, subjective space, and thus as indicating the presence and development of the embodied self. Space precludes a detailed examination of this topic. There are indeed, too many studies which strongly indicate the presence of early infant self-experience to include here.

perceptual responses, such that those which were repeated can become familiar, and those which are presented later can be detected as different. This ability would not be possible unless the various presentations appeared to a single, more-or-less united subjectivity, in which there was the sense that it was the one and same self to which these diverse experiences were presented. It is thus taken by researchers and theorists that infants who are capable of habituation must have the capacity to make the required connections between past and present perceptual responses, and that such infants, therefore, have the sense of self required for such connections to be possible. The frequent use of habituation in infant studies indicates that infants do not experience the world and themselves as an undifferentiated mass, but rather that, through their bodies, they experience themselves strongly and distinctly at a much younger age than Merleau-Ponty's stipulated six months. The testing of infants' early olfactory discriminative abilities bears this out.

6.7.2 Olfactory Discrimination

Using the principles of habituation, the abilities of infants to discriminate between diverse olfactory stimuli have been tested. Olfactory discrimination is only possible if an infant can connect stimuli experienced earlier with stimuli experienced later. In one study, the abilities of three-day old infants to discriminate between the smell of their own mothers' milk, and the smell of milk from other mothers were tested. This test comprised of infants being placed on their backs with breast pads on either side of the head. One of these pads was taken from their own mothers and one was taken from a different mother. The results of the test showed that irrespective of on which side the pads were placed, the infants 'reliably' turned towards the pads taken from their own mothers. This is taken as evidence of infants' ability to 'know' the difference between the odour of their own mothers and the odour taken from elsewhere. In other words, to have the ability to distinguish between the familiar and the unfamiliar. The ability to make this distinction is entirely dependent on the infants' being able to co-ordinate their bodily

experiences in some way, so that connections and comparisons between earlier and later stimuli are possible. The ability to may these connections and comparisons is taken by researchers to indicate that the infants have rudimentary self-awareness, described by Daniel Stern as the 'emergent' self.

In contrast to Merleau-Ponty, Stern argues that the first two months of life are rich with the structure and organisation of a self that 'will remain active for the rest of life.' Stern argues that infants *cannot* be held to experience a lack of organisation. We can only imagine such a lack of organisation because we understand what organisation is like. Stern claims that the infant's first experiences are implicitly differentiated experiences, which become 'yoked' together into a single reference point, causing the infant to experience an initial sense of organisation. As the reference point builds up, the sense of self emerges. Stern claims the self is both the product and process of organisation. He claims that infants' abilities to cohere and amalgamate their experiences, and to develop and expand the body of knowledge we call the self, indicates that the principle of unity is operative right from the beginnings of life (Stern 1985), pp 39- 46.

Stern's concept of the emergent self coheres with the idea of the embodied self as the dynamic unity of experience and ability which develops in tandem with the activities of young infants, and which continues to develop and expand throughout life. This view of the self contrasts with that of the psychological continuity criterion, which takes for granted the ability of humans to have conceptualised experiences, and sees no connection between this ability and the earlier developed bodily organisation which marks the development of young infants. This produces an account of personal identity which alienates the body from experience and mental life, and results in an inadequate appreciation of the body's crucial role in mental life and correspondingly, in personal identity.

6.7.3 Phantom Limbs

Studies of aplasic phantom limbs also supports the case that self-awareness is first experienced through the body.[146] There are strong indications that phantom limbs are associated with the body image or the body schema, or even with both. The body schema and the body image are key components of the embodied self, as they relate to instinctive knowledge about our body and body parts. Recall that the body schema is the most basic form of body-knowledge. It comprises a set of 'motor functions' which operate without self-conscious or intentional self-reference, and involves the body's unreflected posture and movements. By contrast, the body-image is intentional body-knowledge, and involves perceptual, conceptual, and emotional responses. Normally, these two forms of body-knowledge are free to operate separately or together as required.

We have, for example, basic non-reflective knowledge of where body parts are when conducting activities such as walking. In normal walking, only the body schema is used. But when steps are consciously directed, such as when avoiding puddles, the body image is also used. Cases in which the body schema or the body image are damaged are instructive in detecting the presence or absence of bodily self-awareness. Neural disorders such as *hemi-neglect* and *deafferention* are examples of damaged body schemas and body images.

Hemi-neglect occurs when the body schema is present, but the body image is absent. In one reported case of hemi-neglect, a patient has a persistent perceptual defect, relating to the left side of her body. She neglects that side, forgets to dress it, or to comb her hair on that side, and so on. Yet, this patient walks and moves the neglected side normally, using hands and fingers in the usual way. Her instinctive actions are unaffected, but her reflective ones are. Deafferention occurs when the body image is damaged, but the body schema is

[146]'Aplasia' is the congenital absence of a limb, that is, a limb which is absent from birth.

still present. In one case of deafferention, a patient lacks proprioceptive and tactile input from the neck down. Control of limbs below this level thus does not occur in the normal non-reflective way. The patient must therefore exert conscious cognitive and visual control, much as we would with objects external to ourselves. In both of the above cases, the body schema and the body image are disconnected (Gallagher and Meltzoff 1996), pp 214-216.

Similarly, a disconnection between the body schema and the body image is thought to occur when phantom limbs are experienced. The development of phantom limbs in cases other than aplasia or early amputations is understood to indicate a conflict between the body schema, in which the limb is still 'present', and the body image, in which the limb is absent. Gallagher cites that prior to 1961, the fact that aplasia and amputations prior to age six did not produce phantoms was taken to indicate that the body schema had developed since birth, and thus did not include the absent limb. It was assumed that, because there was no conflict between schema and image, a phantom was not produced. But the development of more accurate study techniques has caused these earlier findings to be questioned (Gallagher and Meltzoff 1996), pp 217-218.

In one post 1961 study of thirty aplasics, 17% were found to experience phantom limbs. These phantoms usually appear between the age of five and eight years, but can appear later when under stress, or when undergoing life-changes such as puberty. According to Gallagher, these findings indicate that earlier views that aplasic phantoms are not part of the body schema, and that the body schema is not present at birth, require reconsideration. For, unless the body schema, in which all limbs were present, existed at or prior to birth, the appearance of a phantom limb in the place where a limb had never existed required explanation. Gallagher suggests that the fact aplasics had not experienced phantoms from birth did not automatically mean that the body schema was absent at birth, but

could mean that the early experiences of phantoms had been forgotten, or that the phantoms were present, but, because of immaturity, were inadequately apprehended. While these findings does not *prove* that the body schema is innate, they strongly to support that view and are certainly not inconsistent with it. Gallagher cites other studies[147] which indicate that the body schema is innate, but is open to modification 'by multimodal sensory experiences' throughout life (Gallagher and Meltzoff 1996), pp 214-219.

In addition, Gallagher suggests that revised testing techniques indicate that the body image might also be innate. Earlier theorists had presented confused findings regarding whether phantoms belonged to the body schema or the body image. One example was Simmel, who appropriated the phantom to the schema, but described it in terms of the image, that is, as a conscious representation, in which particular feelings were experienced in the phantom, such as a pain or an itch. Such experiences would indicate that the phantom was experienced explicitly, and was therefore part of the body image, rather than the body schema. Gallagher holds that these findings indicate that the body image could be present in some rudimentary form at birth, but, due to the difficulty in distinguishing between the body schema and the body image, this should not be taken to indicate that the body schema is not. In fact (as outlined below), Gallagher claims that studies of infant imitative abilities add weight to the view that the body schema is innate also (Gallagher and Meltzoff 1996), pp 217-220.

While aplasic phantom limb studies do not yet answer all the questions concerning the body schema and the body image, they do indicate that aplasic phantom limbs are connected in some way to an early form bodily self-awareness which appears to be innate. And while the body schema and the body image are still not perfectly understood, what we can know of them adds considerable weight

[147]For example: Melzack, 1989; Poeck, 1963; Scatena, 1990; Vetter & Weinstein,1967; and Weinstein *et al.*,1964 (Gallagher and Meltzoff 1996), p 214.

to the case that a primitive bodily self-awareness is present early in life. This bodily self-awareness is integral to early infant activity, and is thus the natural precursor to the later developed cognitive awareness. The connection between bodily self-awareness and cognitive self-awareness is not taken into account by Parfit and other psychological continuity theorists, as their focus is on mental life itself, rather than on the preconditions which permit mental life to exist. This results in a misunderstanding of what mental life involves, and correspondingly, in what personal identity involves also.

6.7.4 Imitation

The ability of young infants to imitate simple bodily gestures is taken by many researchers to indicate that a bodily sense of self is present. Such gestures could be facial expressions or tongue protrusions, and are thought to be possible only due to an instinctive 'recognition' between one's own body parts and the body parts of others. This would not be possible without a sense of one's body, and without a rudimentary sense of unity between one's body parts. Also important to imitation is the ability to distinguish between oneself and the object or person imitated. Acts of imitation indicate the presence of Neisser's interpersonal self. Through these acts, selves respond to other selves or persons. I relate to you, I want to be like you, so I copy you. Clearly, one cannot respond to others in this way (or indeed, in any other way) if one does not have a sense of oneself as a self which is separate from others. The apparently instinctive imitation abilities of young infants thus indicates their sense of bodily integration and their sense of being a self distinct from other selves.

Gallagher's studies of infant imitative abilities is part of his exploration of the body schema and the body image. Piaget's view had been that a child under eight months lacked the relevant intellectual mechanism for imitation. It was also understood that imitation would be impossible without a body schema. Thus, the

(presumed) inability of infants to imitate was taken as evidence of a lack of body schema. Gallagher agrees that imitative abilities require a body schema, but claims that as there is ample evidence of infant's imitative abilities, there is also evidence of an innate body schema. Gallagher cites experiments to support this view. In a series of four experiments, in which newborns and slightly older infants were exposed to a variety of imitative situations, the infants demonstrated an immediate ability to imitate. Each experiment consisted of babies being confronted by an adult engaging in one or more facial gestures, including tongue-protrusion, mouth opening, and head turning (Gallagher and Meltzoff 1996), pp 220-221.

In the first experiment, forty normal newborns, aged between one and seventy-one hours, were each confronted for set intervals, first with mouth-opening, and second, with tongue protrusion gestures. The results of this experiment showed that the infants had a strong tendency to imitate, including the youngest, at only forty-two minutes old! Subsequent experiments involved progressively more complex variations in the gestures and included older infants. The data from these studies further supports the arguments for infant imitative abilities, and in the case of the 'older' infants (16-21 days and 6 weeks), an ability to both remember and improve their earlier gestures, and to distinguish between familiar and novel gestures. Commenting on Merleau-Ponty's view that imitation requires a developed body schema, Gallagher argues that the positive response of newborns to imitation clearly indicates that a primitive body schema is present from the earliest moments of life, and thus rejects Merleau-Ponty's view of initial infant indifferentiation (Gallagher and Meltzoff 1996), pp 221-223. Gallagher's and similar findings concerning the imitative abilities of newborns and infants add weight to the case for an innate body schema, and to the view that even the earliest activities of young infants involve a bodily sense of self.

The above studies of infant activities, such as movement, habituation, perceptual discrimination, imitation, and so on, are examples of activities in which the unity of the embodied self is a crucial component. For, as has been explicated, it is only by experiencing my body as *my* body, and as a more-or-less integrated unity under my control, that I can initiate and perform the activities concerned. In demonstrating the role of this early form of self-unity, these studies furnish explanations concerning the bodily unity that is essential to experience and embodied action, which is missing from Parfit's and similar psychological continuity theorists' conceptions of the self. These conceptions do not take seriously the self's role in activity, and consequently, fail to appreciate the self's role in mental life. The relation between embodied experience and cognitive ability becomes more apparent as the development from the immediate experience of bodily self-awareness to the more reflective, mediated self-awareness is explored, in this instance, by considering the formation of autobiographical memories.

6.8 Remembering Oneself

The difference between elementary bodily self-experience and later developed cognitive self-awareness is the difference between direct and reflective experience. Developed cognitive self-awareness involves a self-conscious self-awareness, as opposed to an instinctive or unmediated self-awareness. The transition from one to the other is still inadequately understood, although much has been learnt. The ability to form and access certain types of memories is taken as one of the hallmarks of cognitive development. In particular, the formation of autobiographical memories is important to the development of the cognitive self. Discovering when such memories begin indicates when children first gain a sense of diachronic identity, that is, the sense of being the same person in the present that they were in the past, and that they will be in the future. Understanding the transition from unmediated self-experience to mediated self-experience is helpful to appreciating the role of the body in self-awareness in both instances.

6.8.1 The Beginnings of Self-Identity

Experiments with children have been conducted to discover at what age cognitive self-awareness begins. An important element in these experiments is the discovery of when children are able to identify themselves as being the same persons at a later time that they were at an earlier time. Bodily identity is a crucial factor in this regard, as children's sense of themselves is very much tied to their bodily experiences. In one such experiment, children were videoed, and their ability to recognise themselves was subsequently tested.

This experiment comprised an initial session in which a group of three to five year olds were video-recorded playing a game, and a follow-up session which was conducted a week later. On the first occasion, a sticker was, unbeknown to the child, placed on each child's head during the playing of an 'unusual game.' A week later, the procedure was repeated when the children were playing a different game, in a different location. Three minutes after being videoed, half the children were shown the recent video-recording, while the other half were shown the one taken a week earlier (Butterworth 1998), pp 136-137.

The results from the study showed the that responses to the recordings were different between the older children and the younger children. Whereas the older children shown the recent video responded differently to the older children shown the earlier one, the response between the two younger age-groups was much the same for both videos. More precisely, four to five year olds who viewed the recent recording demonstrated a continuing sense of self with the self of a few minutes earlier. This was evident from reference to their video image as 'me,' and from their attempts to remove the sticker from their own heads. This same age-group who were shown the week-old video-recording responded differently. Apparently aware of the time discrepancy between *then* and *now,* even though they identified themselves with the earlier images, no attempt was made by the

children to remove the sticker. In the case of the three year olds, however, there was no apparent difference in response to the earlier and later videos. In both groups, children attempted to remove stickers, indicating their identification with the video image, but an apparent unawareness of the difference between the self *then*, and the self *now.* Researchers infer from this that the child's sense of self over time, that is, the autobiographical sense of self, emerges at around four years old. This development requires the child's awareness of 'perspective duplication,' otherwise understood as knowing the difference between an earlier and a later self (Butterworth 1998), pp 137-138.

Researchers concluded that children's ability from about four years old to identify with the earlier self implies that their self-awareness is less reliant on visual and proprioceptive feed-back of present perceptual input, than is that of younger children. Butterworth points out that development of the autobiographical self is not simply a cognitive achievement, but is also the ability to 'duplicate perspectives' - to relate the self then to the self now. While researchers could not completely determine the mechanisms involved, the ability could be related to the capacity to represent symbolically, such as through language. Also believed to occur around the same time is the child's ability to attribute mental life to others, sometimes described as the development of a 'theory of mind.' Together, these achievements indicate that the mental self is a symbolic self, while the ecological and interpersonal selves are non-symbolic (Butterworth 1998), pp 138-139.

These experiments indicate that the self is the outcome of a complex developmental process, whose earliest phases involve proprioceptive responses. These responses are initiated by bodily and social engagement between subject and environment, resulting in a 'core' sense of self. The mental self is secondary to, and develops out the 'core' self, and is thus just one aspect of a 'many-faceted' self. While this self ultimately entails a cognitive, remembering, introspective self, it only does so in virtue of its initial proprioceptive, bodily engagement with the

environment and with other selves (Butterworth 1998), p 139. These findings concerning the transition from the earlier, bodily self-experience to the later, cognitive self-experience demonstrate that the embodied self is an important precursor to mental life, and thus furnish explanations as to how that mental life becomes possible, which are missing from the psychological continuity criterion's conception of the self.

6.8.2 The Emergence of the Cognitive Self

The relation between the cognitive self and the development of autobiographical memory has been further explored by some researchers. An issue of concern to these researchers is the occurrence of apparent 'infant amnesia,' that is, the inability of adults to remember anything prior to the age of two years. It seems incongruous that infants display considerable early memory ability, yet none of this earlier period can be recalled later. The period from which adult memories are first recalled coincides with the onset of the cognitive self. It had been held that the cognitive self requires certain neural structures for memory experience, yet infants appear to experience memories much earlier than when these structures develop.

Research by Howe and Courage indicates that the emergence of the cognitive self coincides with the emergence of ability to verbalise, and consequently, to integrate the contents of autobiographical memory, rather than with the ability to experience the memories themselves. Howe and Courage maintain that the mechanisms required for memory are operative much earlier than when memory is first evident. Memory requires 'hardware' and 'software.' Hardware refers to the necessary neurological structures, software the skills, knowledge, and language needed for memories to form. Howe and Courage argue that both 'structures' are operative early in life, even for neonates.[148] They contribute to of all types of

[148]'Neonates' are infants during the first four weeks of life (Tortora 1980), p 625.

memory formation, including autobiographical memory. Apparent changes which occur when memory ability is first evident is not due to sudden onset of previously unavailable skills, but to the continuing development of 'software,' specifically of the cognitive self. Memory does not emerge later as a separate system, discontinuous with previously operating neural processes, but is rather part of a single, dynamic, ongoing process (Howe and Courage 1997), pp 499-500.

An important aspect of Howe and Courage's research is their distinction between the *development* and the *emergence* of the cognitive self. The cognitive self is the culmination of a developmental process which begins very early in the life of human infants. Habituation and early recognition responses indicate the presence of much neural processing during the earliest stages of infant development. Although this early activity is taken as to indicate an emerging sense of self, the immaturity of infants' neural processing mechanisms prevents the emergence of the cognitive self until some time later. Initially, the human infant undergoes a steady process of maturation, until about two years of age, when an important developmental shift occurs. The shift involves infants reaching a certain stage of neural development.

Howe and Courage argue that the level of neural change involved does not match the change in memory ability manifest at the same time. Increased memory ability is not, as was previously held, due to the sudden maturation of the hippocampus, but rather, to the development of certain neocortical connections. They argue that the hippocampus in fact, develops much earlier than previously thought. Rather than a sudden change, development occurs steadily throughout. As the relevant brain structures mature and develop, so also does the capacity for learning and categorising information. This increased capacity leads to a change in the framework through which infants experience the world, permitting the emergence of what has already been developing for some time, namely, the cognitive self. A

major result is the end of infantile amnesia, and the commencement of autobiographical and other forms of event memory. Howe and Courage argue that contrary to earlier views that language and social factors govern the shift in the type of memory experienced, language is merely 'ancillary' rather than 'deterministic' of memory development. They argue that if language was the decisive factor, memories would not appear until much later than they do, such as in the immediate pre-school years. The researchers conclude that the development of language skills is not the primary impetus for development of the cognitive self, but rather, the result of it (Howe and Courage 1997), pp 500-504.

The emergence of the cognitive self heralds the commencement of the infant's ability to verbalise memories. While memory ability does improve, it is out of proportion to the accompanying language development. The apparent jump in memory skills is more likely to be due to improved verbal and narrative skills. This is indicated by studies in which children were questioned concerning past events in their lives. Questions were wide ranging, covering events from the previous day to a year earlier. Some questions concerned periods in the child's life *before verbal skills were acquired.* In many instances, the children's recall of earlier events was 'highly accurate.' indicating that these children were accessing memories previously retained in a non-verbal form. These results support the presence of early infant memory capacity, and correspondingly, of early infant self-awareness. The onset of autobiographical memory thus cannot be seen as an isolated developmental phenomenon, but rather as the outcome of an already occurring intricate and fertile process, which, in the first instance, involves bodily engagement and bodily self-experience (Howe and Courage 1997), p 505. The importance of this early bodily self-experience to the development of cognitive awareness is not recognised by the psychological continuity criterion, and it is thus unable to account for some of the most important factors involved in the possibility of mental life and experience. This results in the account of personal

identity, furnished by the psychological continuity criterion, being inadequate in some of its most crucial aspects.

6.8.3 Autobiographical Memories

Autobiographical memories were described earlier as having special relevance to the self.[149] This relevance persists throughout life, as well as when autobiographical memories first appear. Brewer defines autobiographical memory as memory which contains information relating to the self. The self is characterised as complex, comprising (at least) the *ego* and the *self-schema*. The ego is the subject of experience. It is conscious, has first-person awareness, is phenomenologically aware, and is subject to space and time. The self-schema is a 'cognitive structure,' and consist of self-knowledge. It is formed from repeated instances of ego-self memories. These accumulate to form generic self-memories. The self-schema changes only gradually, thus providing a stable background to the self. Together with relevant aspects of autobiographical memory, the ego and self-schema form the unity of the self. Brewer characterises the *individual* as a more comprehensive construct, comprising self, mind and body, and various skills, such as cognitive or motor skills. The individual includes non-personal aspects of the self, such as general knowledge or other non-personal knowledge (Brewer 1986), pp 26-27. As I conceive the self to be a dynamic unity of all the elements of experience, I find Brewer's self overly mental, and would, therefore, include Brewer's 'individual' as part of the dynamic self.

Autobiographical memories are those memories that comprise life the events which are significant to the sense of ourselves and our lives. They are memories through which we apprehend the past, and gain insight into the future. Autobiographical memories concern persons' lives, persons' views of themselves and of their actions. These memories concern first-person and third-person

[149]Autobiographical memories were referred to in Chapter 4.6.

memory knowledge and have intra-subjective and inter-subjective aspects. Autobiographical memories are more than chains of events. They are part of the framework of the self, and are ways in which selves become known.

Neisser suggests that autobiographical memory can be understood in spatial terms. Drawing on Gibson's work in ecological psychology, Neisser suggests a nested structure, in which smaller units are encased in larger units. For example, sitting in my ergonomic chair means also being in my study, in my house, in a certain street, in a certain town, and so on. To refer to me at this particular place means locating me within this spatial structure. Similarly, a particular memory, say, my attending my daughter's debating finals last week, is situated inside other memories, such as what else I did that week, what happened that day, earlier debates I had attended, and so on. Each item of experience can be seen in this way, as part of a sequence which is in turn part of a larger sequence. Like a set of Russian dolls, each life-event is situated within another event, which, due to its proximity, gives the event shape, dimension, context, and meaning. This approach avoids characterising autobiographical memory atomistically, or merely serially, respecting instead its essentially holistic nature (Neisser 1986), pp 72-77.

Understood holistically, autobiographical memories contribute to our self-formation. We define ourselves in relation to a present which is coherently connected to a past. Our known past is part of what defines us to ourselves. The things we know about ourselves, such as our name, our family, our job, our friends, and so on, all contribute to how we see and know ourselves. In one way or another, all our actions contribute to our autobiography, and hence to our unity, and correspondingly, to our self-formation. Even forgotten activities make their mark. Events unremembered have still contributed to our self-making, and are thus still 'part' of us. There may be subconscious memories that we cannot access. They may not be part of our explicitly self-conscious 'memory-chain,' but

are nevertheless, part of what makes us the selves that we are, and hence, the persons that we are. But as with other aspects of self-development, the psychological continuity criterion fails to take account of these aspects of the self, and thus fails to appreciate self's cumulative aspects. It misses the point that our memories contribute to the development and richness of the self, and that each self thus has its own unique history. This contrasts with the view of the self and memories as being somewhat fleeting, rather superficial phenomena, subject to transfer and manipulation, without any real significance to mental life or personal identity. In taking this view of the self, the psychological continuity criterion fails to account for the sense we have of ourselves over time, as not just persons here and now, but as persons with a particular history, and with a particular sense of ourselves in relation to that history. This sense of ourselves through time is sometimes understood through the concept of 'narrative identity'.

6.8.4 Narrative Identity

Understood in the context of a whole life, autobiographical memories form our life narratives. Persons are mortal creatures. We were born one day, and will die one day. Our lives span the period in between. There are thus boundaries to our life-activities. Realising our mortality, we can gain meaning by looking at our lives as a whole. Life is more than one event merely following another. Lives have shapes, dimensions, and structures. We can reflect on our lives, our reasons, emotions, choices, decisions, and actions. We can remember many of our actions, and sometimes only the event itself. But we do not remember everything. If we did, we would be overloaded with memory, and could probably make little sense of it. A person who remembers everything may as well remember nothing. Memory is essentially discriminating, constrained by many factors. In remembering we typically know who we are, where we have been, where we are heading, and what options lie before us. By visualising the future, we can decide between the available alternatives as to how we should act. To choose, we must be able, in principle at least, to predict the outcome. This requires that we identify the self of

today with the self of the past, and the self of the future. If I suffered from amnesia, and thus lost sense of myself existing in time, such that I had sense of neither my past, nor my future, I would be unable to act. Not knowing what I am doing now, what plans I had in the past, I would not know what choices I should make, or indeed, what choices I *could* make. If I did not remember what my projects were, I could have no intentions to carry out those projects, nor could I see how my present life or my future life connects to my actions. Seeing my life as a narrative grounds my projects, and thus informs me of how I should act now. The things that constitute my narrative are the things that make me the *self* that I am.

Narratives in stories and films often portray events whose significance is not apparent at the time, but which becomes apparent later. The significance of earlier events is often only revealed when married to later events. More often than not, the significance is derivable from the way the event in question fits into a whole pattern of events. Similarly, real lives gain significance from the way their parts relate to each other, especially the most crucial events. Events which endow one's life with significance may not always be known or recognised at the time they occur, sometimes not until after one's death. In this case, only others would know the significance, and not oneself. We can gain some idea of our lives' significance from looking back from periodic vantage-points. We can never look back on our life as a complete whole. We cannot see our lives as a complete narrative, but only as an ongoing one. Seeing our lives as narratives enables us to perceive meaning in our lives as a whole, rather than just in parts. Our narrative at any given time is the unity we have of our self at that time.

There may or may not be meaning in parts of persons' lives, but the meaning of a whole life may not apparent from looking at only the parts. Malpas notes that recounting our life as narratives allows us to look beyond individual events to the

way in which these events are unified into a single life (Malpas 1999), p 80. Having a master narrative gives one's life even more meaning, as it provides a wider setting within which one's life can be shaped and interpreted. Considering life in terms of a narrative structure facilitates the act of self-interpretation. We can step back from our lives to take meaning beyond the present moment, thus expanding the unity of the self. Freeman notes that the impetus to integrate memories can sometimes cause us to confer meaning not intended at the time. Taken too far, this could result in fictions rather than narratives (Freeman 1993), pp 5-11. Integrating false information into the self could lead to our self-concept becoming inaccurate and, therefore, self-deceptive. This appreciation of how our memories, activities, goals, and so on, weave into our life-narratives is missing from the psychological continuity criterion. By conceiving of the self in superficial terms, it fails to appreciate the self's role in our developing self-concepts, and in the sense we have of ourselves as persons with whole lives, rather than as persons with just this experience here and now.

Narrative and time are especially involved in self-formation. Involuntary memories play a special role here, as they refer to things which have occurred, but which have 'congealed' into the self, and are recalled in response to certain triggers which are applied accidentally, rather than intentionally. Proust's *Remembrance of Things Past,* gives many examples of such memories, and demonstrates how these memories are direct and non-interpretive. Often triggered by non-visual senses such as taste, touch or smell, involuntary memories are often ineffable, mysterious, unfinished and lingering. This is because they are directly experienced through the body, and we therefore struggle to find the right cognitive connections and interpretations. For example, Genevieve Lloyd refers to the incident in which Proust's taste of madeleine evokes strong memories of past places and times:

> The essence yielded by involuntary memory is already grasped as a truth which lies not outside but within himself. His seeking mind feels 'overtaken by itself' (Lloyd 1993), p 130.[150]

Involuntary memories like this allow us to revisit and recapture our past selves. Malpas holds that place is important to involuntary memories. Activities are tied to places. As embodied, spatial creatures, place is thus integral to our selves and our memories. Revisiting places can draw threads of the past together, invoking our memories, causing us to reincorporate these memories into our present selves (Malpas 1999), pp 182-184. Our self-knowledge is widened and the integration between past and present aspects of the self is strengthened. This aspect of our selves is missed by the psychological continuity criterion, which, while failing to recognise the self's importance to memory, also fails to recognise the importance of memory to the self.

Voluntary memory can be used deliberately in the act of self-formation. Autobiographies are often written with this aim. The earliest known autobiography, Augustine's *Confessions,* employs voluntary memory to incorporate the past *prior* to his conversion, with the awareness gained *since* conversion. In doing so, Augustine transforms the past, and sees it, rather than just a segment of life, but as part of a single narrative. In the act of narrative, he steps back from himself, viewing past and present self-concepts in light of his whole life, and in light of his

---

[150]Proust 'recalls' an incident in which, when having afternoon tea with his mother, memories are generated by the taste of the madeleine (small cake), dipped in tea: 'Many years had elapsed during which nothing of Combray, save what was comprised in the theatre and the drama of my going to bed there, had any existence for me, when one day in winter, on my return home, my mother, seeing that I was cold, offered me some tea, a thing I did not ordinarily take. I declined at first, and then, for no particular reason, changed my mind. She sent for one of those squat, plump little cakes called "petites madeleines", which look as though they had been moulded in the fluted valve of a scallop shell. And soon, mechanically, dispirited after a dreary day with the prospect of a depressing morrow, I raised to my lips a spoonful of the tea in which I had soaked a morsel of the cake. No sooner had the warm liquid mixed with the crumbs touched my palate than a shudder ran through me and I stopped, intent upon the extraordinary thing that was happening to me. An exquisite pleasure invaded my senses, something isolated, detached, with no suggestion of its origin. And once the vicissitudes of life had become indifferent to me, its disasters innocuous, its brevity illusory - this new sensation having had on me the effect which love has of filling me with a precious essence; or rather this essence was not in me, it *was* me' (Proust 1981), p 48. Because Proust's memory was predominantly sensual, that is, bodily experienced, it was not immediately or easily captured cognitively.

understanding of his life as related to God (Lloyd 1993), pp 14-42; (Augustine 1996). In widening his self concept in this way, Augustine demonstrates that the self is not a static thing, but is a dynamic unity, which is subject to constant change and 'updating.'

Narratives can unify the self by means of underlying principles or goals. These principles can be the primary motivation for actions, unifying a person's self and a person's life simultaneously. For example, a person whose primary aim is to pursue a particular sport may judge all his activities in relation to that sport. The choice of whether to perform one activity or an alternative activity is made on the basis of its contribution to the enhancement of the sport concerned. Individual activities are formed into groups, which, united with other groups, form the person's life-history. Similarly, argues MacIntyre, life narratives can be understood as divisible into episodes. Episodes comprise segments and segments comprise individual actions. According to MacIntyre, a life underpinned by virtue is united by the value-judgments which guide that life's actions. Even when judgments are hidden, they are the fundamental principles which draw that life into a single history. Understanding a single action on its own may make no sense. Only by conceiving it as part of a whole history can its meaning be understood. MacIntyre argues that a fragmented, temporary self, such as that of Parfit, cannot incorporate values into a single life in this way. Fragmented selves do not last a whole life-time, and thus cannot be underpinned by sets of values which require a life-time to mature (MacIntyre 1981), pp 190-209. Thus, selves which incorporate long-term values into their unity are likely to be more enduring selves that those which do not.

The concept of narrative draws on the self's essentially dynamic nature. Defined by its actions and memories, the self is necessarily vital and active. Paul Ricoeur holds that narrative more fruitfully accounts for the self's durability amidst change,

than Parfit's 'science fiction.' Ricoeur argues that 'literary fiction' more aptly represents the real lives from which it is taken, than 'unrealizable' thought experiments. Literary characters are not static, yet when changed by the events and plots of their stories, they are not assumed to have changed identity:

> According to my thesis, the narrative constructs the durable character of an individual, which one can call his or her narrative identity, in constructing the sort of dynamic identity proper to the plot (*l'intrigue*) which creates the identity of the protagonist in the story. It is primarily in the plot therefore that we must search for the mediation between permanence and change, before being able to transfer it to the character. The advantage of this detour through the plot is that it furnishes the model of discordant concordance upon which it is possible to construct the narrative identity of a character. The narrative identity of this character will only be known correlative to the discordant concordance of the story itself (Ricoeur 1991), pp 77-78.

A character's identity is maintained against impermanent and changed surroundings. Retention of identity is compatible with change, not contrary to it. Unless a character could change and grow, it is difficult to imagine how there could be a story at all. Replicas, claims Ricoeur, do not interact with other selves, and so are not changed by them. 'Narrative situations' inform of real lives, and can act as models by which to interpret lives. This interpretation sustains our selves, permitting growth and stability amidst change (Ricoeur 1991), pp 77-79. In narrating our lives to our selves, we develop and progress, while at the same time, drawing together the diverse parts of our lives into the unity of our selves. These dimensions of the self are overlooked by the psychological continuity criterion. For Parfit and similar theorists, the self is superficial and insubstantial. But as has been argued, this view of the self is mistaken, and results from incorrect preconceptions about what the self is, and why the self is integral to experience and mental life. When the various forms of self-unity are more closely examined, it becomes evident that self-unity is crucial to all forms of experience

and mental life, such that it is *only* in view of the self's unity that experience and mental life are possible.

6.9 Summary

This chapter has addressed the psychological continuity criterion's Reductionist conception of the self. This conception holds that selves do not ground identity, and cannot be presumed to exist. On this view, a person's identity is theoretically transferable, indeterminate, and ultimately insignificant. I have argued that the Reductionist conception of the self has little basis and leaves subjectivity inadequately explained. In contrast to the Reductionist view, I have argued that the self is essential to experience, and that the self is given in experience. I have also argued that the self is a dynamic unity, which is first experienced bodily, and is later experienced cognitively. In support of my claims, I have presented philosophical arguments from Kant and Malpas, and considerable evidence from empirical studies. I have also explored the relevance to the self of autobiography and narrative.

To facilitate understanding the self, I have discussed the self in terms of the cognitive self and the embodied self, as these distinctions permit exploration of different aspects of the self. Drawing on Kant and Malpas, the cognitive self was shown to be an essential component of mental life and experience, as this aspect of the self is understood to be the apperceptive unity of mental life and experience. I have argued that it is only by the integration of the various aspects of mental life and experience into a single subjectivity, that experience and mental life are possible. In support of the embodied self, I have referred to studies of early infant development, which showed various ways in which the self is bodily experienced early in life, and how this bodily self-experience is essential to, and an essential part of, infants' early abilities and activities, such as motility, habituation, imitation, and so on.

To explore ways in which the embodied self develops into, and, in some sense, gives way to, the cognitive self, aspects of concept-formation were examined. Studies addressed included those concerning infants' ability for self-identification, and for the formation of autobiographical memories. These abilities mark the onset of concept-formation, and the subsequently, the capacity for self-awareness and self-reflexivity. Having self-awareness and self-reflexivity permits the unity of self to include our past 'selves' and our anticipated future 'selves.' The concept of narrative was also examined, to elucidate various ways in which self-knowledge can expand and develop, and in which whole lives can gain coherence and meaning. Narrative permits us to imbue our lives with value, and to view our lives as coherent wholes, and thus provides a richer conception of the self than a conception which conceives of experience as disconnected and atomistic. In summary, these arguments in support of the self demonstrate several important ways in which the self is a unity, which is both a part of experience, and a necessary condition of experience. These arguments also show that the self which is conceived as a dynamic unity is not a fleeting, ephemeral entity, but is a self which is structured, coherent, organised, inherently complex, and, importantly, enduring.

Chapter 7 Conclusion

7.1 Central Claims

The foregoing chapters have attempted to defend this thesis' two central claims. These claims concern the psychological continuity's atomistic approach to personal identity, and the proposal of an alternative, more holistic approach to this issue. Specifically, as outlined earlier, the central claims of this thesis are:

- That due to its reductionist, criterial approach, the psychological continuity criterion is inadequate to account for personal identity.
- That a sound approach to personal identity must respect the complex, dynamic, holistic, non-reductive nature of persons.

In respect of the first claim, I have argued that by selectively categorising personal identity in terms of linear mental chains, the psychological continuity criterion produces a misleading account of persons, and consequently, of personal identity. Because it neglects many other identity-determining features, the psychological continuity criterion produces an untenably thin account of the conditions under which personal identity is maintained. The continuance of event memories, certain beliefs and intentions, and so on, without the specification of what this continuance involves, results in 'persons' who are unfamiliar to us. Their minds are theoretically isolated from the world in which they live, their mental contents disordered, their body's irrelevant, and their self-knowledge superficial. Both their individuality and their long-term survival are distorted. As a result, their personal autonomy and their capacity for responsibility are distorted also. My concern is that these altered conceptions of persons are detrimental, both to individuals and to societies. Such persons are anomalous, and do not represent persons as they are, or as we know them to be.

As the foregoing chapters show, these aberrant outcomes result from an analysis which is inappropriate due to several deficiencies: undue focus on thought

experiments which concern 'uncommon' persons, to the detriment of real persons (Chapter 2); inadequate appreciation of the extent to which the external world and our perception of the external world influences and constrains our mental lives (Chapter 3); inaccurate characterisation of the mind's structure and the failure to appreciate the relevance of the mind's particular ownership (Chapter 4); neglect of the body's role in mental life (Chapter 5), and finally, confusion with the role and nature of the self and its essential place in mental life and bodily activity. Together, these misapprehensions of personal identity factors yield an account of personal identity which is metaphysically and ethically unsatisfactory, thus supporting the first claim: that due to its reductionist, criterial approach, the psychological continuity criterion is inadequate to account for personal identity.

In respect of the second claim, I have throughout the foregoing chapters advanced two major points. First, personal identity involves essential features not specified by the psychological continuity criterion. Second, due to their inherent interrelatedness, many of these features, although concealed, are nevertheless implicated by the psychological continuity criterion, but are entirely ignored by it. Because they are related at a fundamental level, these features are misrepresented when conceptualised in isolation from each other. They are missed by the superficial analysis of thought experiments (Chapter 2), and include at least the external world (Chapter 3), the individual and unique ownership of mental states (Chapter 4), the body (Chapter 5), and the dynamic unity of the self (Chapter 6). I have argued that being a person, and having a mind necessarily involves living in, and experiencing a world, having a coherently structured mind, having a body, and experiencing oneself as a more-or-less united self. As I have also argued, the mind's holism is related to our status as acting as well as thinking beings. From the earliest moments of life, to the last moments of self-reflection, persons are (in principle) interrelated, active beings. Their mental lives and activities depend on many parts, and on the relations between these parts.

Removing parts, such as bodies and environments, from persons makes no sense, as these things are part of what persons are, such that if these parts were changed, persons' identities would be changed also. These arguments, which have been elucidated and defended throughout the thesis, support my second claim: that a sound approach to personal identity must respect the dynamic, holistic, non-reductive nature of persons.

In light of these and the above arguments, it is clear that an account of personal identity which does not address the wealth of empirical evidence, available from *real* persons in *real* situations, is an account which has no authentic foundation, and which, therefore, cannot be sustained. The very fact that personal identity is about real persons vindicates the merits of the empirical evidence used throughout this thesis. As has been shown, pure speculation is simply inadequate to answer the issues which personal identity involves. Empirical evidence is particularly important to the psychological continuity criterial conception of personal identity because, as mentioned earlier, its disregard for bodily identity is based on speculation alone, and does not take into account the issues raised throughout this thesis (see the last paragraph in Section 1.3, including footnote 11).

The above conception of persons now needs to be considered in the light of concerns expressed earlier in this thesis.[151] Recall that 'Parfitian Persons' are 'series persons,' whose endurance over extended periods of time is dubious. The fragmentary nature of series persons means that they have reduced responsibility for the actions of earlier series persons. My concern is that this reduced conception of personhood, and the diminished conception of responsibility which goes with it will lead to less respect for others, and to less appreciation of the implications of one's actions. On my account of personal identity, however,

[151]See Chapter 1.5.

persons are construed to be enduring entities, which are not destroyed or threatened by change, but which, on the contrary, thrive and develop through change. Thus, change *per se* would not prevent persons from bearing responsibility for their actions. This is not to say that change would not be a factor in, for example, bearing responsibility for crimes. Clearly, there are many instances where change would be important, such as in the case of remorse, old age, illness, and so on. But the difference between my account Parfit's is that the *type* of change would be relevant, rather than the *amount* of change. Because persons consider themselves as whole entities, change could mean reform, reparation, self-improvement, the development of wisdom, and so on, rather than fragmentation, separation, and ultimately the death of one's former self. Contra Parfit, I consider that this conception of persons is more personally liberating, and more advantageous to society as a whole.

7.2 Future Directions

While the two major claims of this thesis, and the accompanying subordinate claims, are not intended to provide a definitive account of personal identity, they are intended to indicate some of the features which such an account should incorporate. A more developed account of personal identity could be produced by undertaking further research into additional areas, such as the cultural, economic, and social forces which influence our lives. These forces are part of the world in which we live, and thus, on the basis of this thesis, are contributors to our identities. As has been argued throughout this thesis, our identities are made of many things. Knowing more about what these things are, and how these things affect us increases our scope for self-knowledge and self-improvement. Taking the factors which compose personal identity seriously is thus more personally and socially useful than is taking those factors for granted. But while there is much more to be learnt about personal identity than can be covered in this thesis, it is my view that because persons are such complex, dynamic entities, the most

extensive research will never reveal all there is to know about persons or personal identity. All the same, an account of personal identity which recognises our inherently dynamic nature, and which acknowledges that the world in which we live is intimately related to our thoughts and actions, gives more insight into personal identity, than an account which fails to acknowledge the sources and influences from which those thoughts and actions first arise.

Bibliography

Allan, T., C. Phillips, et al. (1999). Spirits of the Snow. Amsterdam, Time-Life Books BV.

Allen, D., Ed. (1997). Culture and Self. Boulder, Colorado, Westview Press.

Alston, W. P. (1998). Internalism and Externalism in Epistemology. Routledge Encyclopedia of Philosophy. London, Routledge. 821-826.

Anchor, R. (1995). "Review Essay: 'Michel Foucault: Subversions of the Subject' by Philip Barker." History and Theory **34**(1): 122-132.

Anderson, P. S. (1992). "A Question of Personal Identity." Personalist Forum **8**(1): 55-68.

Anscombe, G. E. M. (1975). The First person. Mind and Language. Oxford, Clarendon Press. 45-65.

Aristotle (1963). The Philosophy of Aristotle. New York, New American Library.

Aristotle (1963). Psychology. The Philosophy of Aristotle. New York, N Y, New American Library. 280-378.

Armstrong, D. (1968). The Materialist Theory of Mind. London, Routledge & Kegan Paul.

Atkinson, J. and O. Braddick (1989). Development of Basic Visual Functions. Infant Development. Hove, UK, Lawrence Erlbaum Associates. 7-41.

Augustine, S. (1996). The Confessions of Saint Augustine. Springdale, PA, Whitaker House.

Ayer, A. J. (1963). The Concept of a Person. The Concept of a Person and Other Essays. New York, St. Martin's Press. 82-128.

Backman, L., Ed. (1992). Memory Functioning in Dementia. Advances in Psychology. Amsterdam, North-Holland.

Baier, A. C. (1991). A Progress of Sentiments. Cambridge, MA, Harvard University Press.

Baillie, J. (1990). "The Problem of Personal Identity." Cogito **4**(2): 106-112.

Baker, L. R. (1987). Saving Belief. Princetown, N J, Princetown University Press.

Baker, L. R. (1995). Explaining Attitudes. Cambridge, N Y, Cambridge University Press.

Baker, L. R. (1998). "The First-Person Perspective: A Test for Naturalism." American Philosophical Quarterly **35**(4): 327-346.

Barrett, W. (1986). Death of the Soul. Garden City, NY, Anchor Press.

Baxter, D. L. (1987). "A Defense of Hume on Identity Through Time." Hume Studies **13**(November): 323-342.

Bechtel, W. (1988). "Connectionism and Rules and Representation Systems: are they Compatible?" Philosophical Psychology **1**(1): 5-16.

Bechtel, W. (1988). Philosophy of Mind. Hillsdale, N J, Lawrence Erlbaum Assoc.

Bechtel, W. and A. Abrahamsen (1991). Connectionism and the Mind. Cambridge, M A, Blackwell.

Beck, S. (1991). "Finding Ourselves: Personal Identity and the Limits of Possible-World Arguments." South African Journal of Philosophy **10**(1): 1-6.

Berkeley, G. (1965). Principle, Dialogues, and Correspondence (1710/1713). Indianapolis, N Y, Bobbs-Merrill.

Berkeley, I. and F. J. Pelletier (1999). Vagueness. The Cambridge Dictionary of Philosophy. Cambridge, Cambridge University Press. 2nd., ed. 945-947.

Bermudez, J. L. (1994). “The Unity of Apperception in the 'Critique of Pure Reason'.” European Journal of Philosophy **2**(3): 213-340.

Bermúdez, J. L., A. Marcel, et al., Ed. (1995). The Body and the Self. Cambridge, MA, MIT Press.

Bilgrami, A. (1992). Belief and Meaning. Cambridge, MA, Blackwell.

Bilgrami, A. (1995). “Book Reviews: A Theory of Content and Other Essays. Jerry A Fodor. Cambridge: Massachusetts Institute of Technology Press, 1991: Holism: A Shopper's Guide. Jerry A Fodor and Ernest Lepore. Cambridge: Basil Blackwell, 1992.” The Journal of Philosophy **92**(7): 330-344.

Biro, J. (1993). Hume's New Science of the Mind. The Cambridge Companion to Hume. Cambridge, N Y, Cambridge University Press. 33-63.

Blatner, W. (1996). “Existence and Self-Understanding in 'Being and Time'.” Philosophy and Phenomenological Research **56**(1): 97-110.

Block, N. (1998). Holism; Mental and Semantic. Routledge Encyclopedia of Philosophy. London, Routledge. 488-493.

Blum, P. (1985). “Deconstruction and Creation.” Philosophy and Phenomenological Research **46**(2): 293-306.

Bohm, D. (1980). Wholeness and the Implicate Order. London, Routledge & Kegan Paul.

Bole, T. J. (1990). “Zygotes, Souls, Substances, and Persons.” Journal of Medicine and Philosophy **15**(6): 637-652.

Borst, C. V., Ed. (1970). The Mind-Brain Identity Theory. Controversies in Philosophy. London, Macmillan.

Brennan, A. (1988). Conditions of Identity. Oxford, Clarendon Press.

Brennan, A. (1990). Fragmented Selves and the Problem of Ownership. The Aristotelian Society. London, The Aristotelian Society. 143-158.

Brennan, A. A. (1987). “Survival and Importance.” Analysis **47**(1): 225-230.

Brentano, F. (1960). The Distinction between Mental and Physical Phenomena. Realism and the Background of Phenomenology. Illinois, The Free Press of Glencoe. 39-61.

Brentano, F. (1960). Presentation and Judgment Form Two Distinct Fundamental Classes. Realism and the Background of Phenomenology. Illinois, The Free Press of Glencoe. 63-70.

Brewer, B. (1995). Bodily Awareness and the Self. The Body and the Self. Cambridge, MA, MIT Press. 291-309.

Brewer, W. F. (1986). What is Autobiographical Memory? Autobiographical Memory. Cambridge, Cambridge University Press. 25-49.

Brewer, W. F. (1988). Memory for Randomly Sampled Autobiographical Events. Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory. Cambridge, Cambridge University Press. 21-90.

Bricke, J. (1980). Hume's Philosophy of Mind. Edinburgh, Edinburgh University Press.

Broad, C. D. (1925). The Mind and its Place in Nature. London, Routledge & Kegan Paul.

Broad, C. D. (1978). Kant An Introduction. Cambridge, Cambridge University Press.

Brogan, W. (1991). “The Decentered Self: Nietzsche's Transgression of Metaphysical Subjectivity.” The Southern Journal of Philosophy **29**(4): 419-429.

Brook, A. (1998). “Unified Consciousness and the Self.” Journal of Consciousness Studies **5**(5-6): 583-91.

Brown, R. and J. Kulik (1982). Flashbulb Memories. Memory Observed. San Francisco, W H Freeman and Company. 23-40.

Brueckner, A. (1993). "Parfit on What Matters in Survival." Philosophical Studies **70**: 1-22.

Buckout, R. (1982). Eyewitness Testimony. Memory Observed. San Francisco, W H Freeman and Company. 116-125.

Burford, T. (1989). "Person, Identity, and Imagination." Personalist Forum **5**(1): 7-25.

Burge, T. (1991). Individualism. The Nature of Mind. Oxford, Oxford University Press. 536-567.

Burkett, I. (1994). "The Shifting Concept of the Self." History of the Human Sciences **7**(2): 7-28.

Bushnell, D. E. (1993). "Identity, Psychological Continuity, and Rationality." Journal of Philosophical Research **18**: 15-24.

Butler, J. (1867). Two Dissertations:
Dissertation 1: Of Personal Identity. The Analogy of Religion. London, William Tegg. 193-198.

Butterworth, G. (1977). "Visual Proprioception and Postural Stability in Infancy.
A Developmental Study." Perception **6**: 255-262.

Butterworth, G. (1992). "Origins of Self-Perception in Infancy." Psychological Inquiry **3**(2): 103-111.

Butterworth, G. (1995). An Ecological Perspective on the Origins of the Self. The Body and the Self. Cambridge, MA, MIT Press. 87-105.

Butterworth, G. (1998). "A Developmental-Ecological Perspective on Strawson's 'The Self'." Journal of Consciousness Studies **5**(2): 132-140.

Butterworth, G. and M. Harris (1994). Principles of Developmental Psychology. Hove, U K, Lawrence Erlbaum.

Butterworth, G. and C. Henty (1991). "Origins of the Proprioceptive Function of Vision: Visual Control of Posture in one day old domestic Chicks." Perception **20**(381-386):

Campbell, J. (1994). Past, Space, and Self. Cambridge, M A, M I T Press.

Carr, D. (1995). "The Question of the Subject: Heidegger and the Transcendental Tradition." Human Studies **17**: 403-418.

Cartwright, H. M. (1993). "On Two Arguments for the Indeterminacy of Personal Identity." Synthese **95**(2): 241-273.

Casey, E. (1987). Remembering. Indiana UP, Bloomington.

Cassam, Q. (1989). "Kant and Reductionism." Review of Metaphysics (43): 72-106.

Castaneda, H.-N. (1966). "'He': A Study in the Logic of Self-Consciousness." Ratio **111**(2): 130-157.

Castaneda, H.-N. (1969). "On the Phenomeno-Logic of the I." Afken des X1V Internationalen Kongress fur Philosophie, University of Vienna **111**: 260-266.

Casullo, A. (1988). "A Fourth Version of the Bundle Theory." Philosophical Studies **54**(July): 125-139.

Chalmers, D. J. (1996). The Conscious Mind. Oxford, Oxford University Press.

Chappell, V., Ed. (1994). The Cambridge Companion to Locke. Cambridge, Cambridge University Press.

Chaturvedi, V. (1992). "Indeterminateness of the Concept of a Person." Indian Philosophical Quarterly **29**(2): 115-125.

Chisholm, R. M. (1950). The Theory of Appearing. Philosophical Analysis. Ithaca, N Y, Cornell University Press. 102-118.

Chisholm, R. M., Ed. (1960). Realism and the Background of Phenomenology. The Library of Philosophical Movements. Glencoe, Illinois, The Free Press.

Chisholm, R. M. (1969). The Loose and Popular and the Strict and Philosophical Senses of Identity. Perception and Personal Identity. Cleveland, OH, The Press of Case Western Reserve University. 82-106.

Chisholm, R. M. (1969). Reply. Perception and Personal Identity. Cleveland, OH, The Press of Case Reserve University. 128-139.

Chisholm, R. M. (1976). Person and Object. La Salle, ILL, Open Court Publishing Company.

Chisholm, R. M. (1977). Theory of Knowledge. Englewood Cliffs, N J, Prentice-Hall Inc.

Chisholm, R. M. (1981). The First Person. Sussex, Harvester Press.

Chisholm, R. M. (1982). The Foundations of Knowing. Sussex, The Harvester Press.

Clark, R. (1988). “Self Knowledge and Self Consciousness: Thoughts about Oneself.” Topoi **7**(1): 47-55.

Clark, S. R. L. (1991). How Many Selves Make Me? Human Beings. Cambridge, Cambridge University Pres. 213-233.

Coad, L. (1996). “Commentary on "Loopholes, Gaps, and What is Held Fast".” Philosophy, Psychology, and Psychiatry **4**(3).

Cockburn, D., Ed. (1991). Human Beings. Royal Institute of Philosophy Supplement: 29 Supplement to 'Philosophy' 1991. Cambridge, N Y, Cambridge University Press.

Cockburn, R. (1985). “Personal Identity Revisited.” Canadian Journal of Philosophy **15**(3): 379-404.

Colpietro, V. (1990). “The Integral Self: Systematic Illusion or Inescapable Task?” Listening **25**(3): 192-210.

Cooper, D. E. (1997). “Wittgenstein, Heidegger and Humility.” Philosophy **72**(279): 105-123.

Cottingham, J. (1984). Rationalism. London, Paladin.

Curtis, H. (1983). Biology. New York, NY, Worth Publishers Inc.

Curzer, H. (1991). “An Ambiguity in Parfit's Theory of Personal Identity.” Ratio (new series) **4**(1): 16-24.

Damasio, A. (1994). Descartes' Error. London, Papermac.

Davidson, D. (1980). Actions, Reasons, and Causes. Essays on Actions and Events. Oxford, Clarendon Press. 3-19.

Davidson, D. (1980). Mental Events. Essays on Actions and Events. Oxford, Clarendon Press. 207-227.

Davidson, D. (1984). Inquiries into Truth and Interpretation. Oxford, Clarendon Press.

Davidson, D. (1987). The Myth of the Subjective. Relativism: Interpretation and Confrontation. Notre Dame, Notre Dame University Press. 159-172.

Davidson, D. (1991). “Three Varieties of Knowledge.” Royal Institute of Philosophy Supplement (30): 153-166.

Davidson, D. (1995). “The Problem of Objectivity.” Tijdschrift Voor Filosofie **57**(2): 203-220.

Davie, W. E. (1984). “Hume on Perceptions and Persons.” Hume Studies **10**(November): 125-138.

Dennett, D. (1976). Conditions of Personhood. The Identities of Persons. Berkeley, CA, University of California Press. 175-196.

Dennett, D. (1978). Intentional Systems. Brainstorms. Montgomery, Vt., Bradford Books. 3-23.

Dennett, D. (1978). Two Approaches to Mental Images. Brainstorms. Montgomery, Vt., Bradford Books. 174-189.

Dennett, D. (1981). True Believers: The Intentional Strategy and Why it Works. The Intentional Stance. Oxford, Oxford University Press. 13-35.

Dennett, D. (1991). Consciousness Explained. Boston, Little, Brown and Company.

Dennett, D. C. (1996). Kinds of Minds. London, Phoenix.

Descartes, R. (1968). Discourse on Method and the Meditations (1637). London, Penguin.

Devitt, M. and K. Sterelny (1987). A causal theory of reference: names. Language and Reality. Cambridge, M A, MIT Press. 55-66.

Devitt, M. and K. Sterelny (1987). Description theories of reference: names. Language and Reality. Cambridge, M A, MIT Press. 39-54.

Devitt, M. and K. Sterelny (1987). Truth and reference. Language and Reality. Cambridge M A, MIT Press. 15-38.

Doepke, F. (1990). "The Practical Importance of Personal Identity." Logos **11**: 83-91.

Donagan, A. (1990). "Real Human Persons." Logos **11**: 1-16.

Dunning, W. (1993). "Post-Modernism and the Construct of the Divisible Self." British Journal of Aesthetics **33**(2): 132-141.

Edelman, G. M. (1989). The Remembered Present. New York, Basic Books.

Edelman, G. M. (1992). Bright Air, Brilliant Fire. New York, N Y, Basic Books.

Ehring, D. (1995). "Personal Identity and the R-Relation: Reconciliation through Cohabitation?" Australasian Journal of Philosophy **73**(3): 337-346.

Eibl-Eibesfeldt, I. (1975). Ethology: The Biology of Behavior. New York, Holt, Rinehart and Winston Inc.

Eilan, N. (1995). Consciousness and the Self. The Body and the Self. Cambridge, MA, MIT Press. 337-357.

Elliot, R. (1991). "Personal Identity and the Causal Continuity Requirement." The Philosophical Quarterly **41**(162): 55-75.

Evans, G. (1982). Frege. The Varieties of Reference. Oxford, Clarendon Press. 7-41.

Evans, G. (1982). Russell. The Varieties of Reference. Oxford, Clarendon Press. 42-63.

Evans, G. (1982). Self-Identification. The Varieties of Reference. Oxford, Clarendon Press. 205-257.

Evans, G. (1982). The Varieties of Reference. Oxford, Clarendon Press.

Evans, G. (1997). Can there be Vague Objects? Vagueness: A Reader. Cambridge MA, MIT Press. 317.

Fann, K. T. (1969). Wittgenstein's Conception of Philosophy. Oxford, Basil Blackwell.

Fate Norton, D. (1982). David Hume. Princeton, New Jersey, Princeton University Press.

Felice, C. (1987). "Causation and the Self: Reply to Frankel." International Philosophical Quarterly **27**(3): 329-334.

Ferrarin, A. (1994). "Husserl on the Ego and its Eidos." Journal of the History of Philosophy **32**(4): 645-659.

Feyerabend, P. (1975). Against Method. London, Verso.

Flage, D. E. (1990). Personal Identity. David Hume's Theory of Mind. London, Routledge. 131-155.

Flew, A. (1968). Lock and the Problem of Personal Identity. Locke and Berkeley. Garden City, N Y, Anchor Books. 155-178.

Flew, A. (1985). "'Personal Identity and Imagination': One Objection." Philosophy **60**: 123-126.

Flew, A. (1993). "People Themselves, and/or Their Selves?" Philosophy **68**: 546-548.

Fodor, J. (1981). Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. Representations. Cambridge, MA, 225-253.

Fodor, J. (1981). Three Cheers for Propositional Attitudes. Representations. Cambridge, MA, MIT Press. 100-123.

Fodor, J. and Lepore, E. (1992). Holism A Shopper's Guide. Oxford, Blackwell.

Fodor, J. and Z. W. Pylyshyn (1988). "Connectionism and Cognitive Architecture: A Critical Analysis." Cognition **28**: 3-71.

Fodor, J. A. (1983). The Modularity of Mind. Cambridge, MA, MIT Press.

Fogel, A. (1993). Developing through Relationships. Chicago, University of Chicago Press.

Forrester, M. G. (1996). Persons, Animals, and Fetuses. Dordrecht, Kluwer Academic Publishers.

Foster, J. (1991). The Immaterial Self. London, Routledge.

Frankel, L. (1987). "Causation and the Self: Reply to Felice." International Philosophical Quarterly **27**(3): 325-327.

Frankfurt, H. G. (1971). "Freedom of the Will and the Concept of a Person." Journal of Philosophy **68**(1): 5-20.

Franz, B. (1960). Genuine and Fictitious Objects. Realism and the Background of Phenomenology. Illinois, The Free Press of Glencoe. 71-75.

Freeman, M. (1993). Rewriting the Self. London, Routledge.

Frege, G. (1952). On Concept and Object. Translations from the Philosophical Writings of Gottlob Frege. Oxford, Basil Blackwell. 42-55.

Frege, G. (1952). On Sense and Reference. Translations from the Philosophical Writings of Gottlob Frege. Oxford, Basil Blackwell. 56-78.

Furst, C. (1979). Origins of the Mind. Englewood Cliffs, N.J., Prentice-Hall, Inc.

Fuster, J., M (1995). Memory in the Cerebral Cortex. Cambridge, M A, M I T Press.

Gale, R. (1994). "James on Self Identity over Time." The Modern Schoolman **71**(3): 165-189.

Gallagher, S. (1992). "The Theatre of Personal Identity: From Hume to Derrida." Personalist Forum **8**(1): 21-30.

Gallagher, S. and A. N. Meltzoff (1996). "The Earliest Sense of Self and Others: Merleau-Ponty and Recent Developmental Studies." Philosophical Psychology **9**(2): 211-233.

Gallagher, S. and J. Shear (1997). "Models of the Self." Journal of Consciousness Studies **4**(5/6): 399-540.

Gallagher, S. and J. Shear (1998). "Models of the Self (Part 2)." Journal of Consciousness Studies **5**(2): 132-244.

Garfield, J. L. (1990). Foundations of Cognitive Science. New York, Paragon House.

Garfield, J. L. (1993). Western Idealism and its Critics. Amherst, M A, Hampshire College.

Garfield, J. L. (1997). Quine and Sellars: Twilight of the Dogmas and the Death of Modernism. American Philosophers. Hobart, Pyrrho Press. 43-76.

Garrett, B. (1995). "Wittgenstein and the First Person." Australasian Journal of Philosophy **73**(3): 347-355.

Gazzaniga, M (1970). The Bisected Brain. New York, Meredith Corporation.

Gazzaniga, M. S. and J. E. Le Doux (1978). The Integrated Mind. New York, Plenum Press.

Geach, P. and M. Black, Ed. (1952). Translations from the Philosophical Writings of Gottlob Frege. Oxford, Basil Blackwell.

Gibson, J. J. (1977). The Theory of Affordances. Perceiving, Acting, Knowing: Towards an Ecological Psychology. Hillsdale, NJ, Lawrence Erlbaum. 67-82.

Gillett, G. (1986). "Brain Bisection and Personal Identity." Mind **95**(378): 224-229.

Gillett, G. (1986). "Disembodied Persons." Philosophy **61**: 377-386.

Gillett, G. (1990). "The Subject of Experience." Logos: Philosophic Issues in Christian Perspective **11**: 93-109.

Goethals, G. R. and P. R. Solomon (1989). Interdisciplinary Perspectives on the Study of Memory. Memory: Interdisciplinary Approaches. New York, Springer-Verlag. 1-13.

Goodman, M. (1992). "A Sufficient Condition for Personhood." The Personalist Forum **8**(1 (Supp)): 75-81.

Gould, S. (1999). Tumble Turns. Sydney, Harper Collins.

Grave, S. A. (1968). The Mind and its Ideas: Some Problems in the Interpretation of Berkeley. Berkeley's Principles of Human Knowledge. Belmont, California, Wadsworth Publishing Co. 10-23.

Grisez, G. (1990). When Do People Begin? Proceeding of the American Catholic Philosophical Association. Washington, D C, A C P A. 27-47.

Hamlyn, D. W. (1984). Metaphysics. Cambridge, Cambridge University Press.

Harvey, R. L. (1994). Neural Network Principles. Englewood Cliffs, N J, Prentice Hall.

Hayward, A. L. and J. J. Sparkes (1988). The Concise English Dictionary. Poole, Dorset, New Orchard Editions.

Hegel, G. W. (1910). The Phenomenology of Mind. London, Allen & Unwin.

Heidegger, M. (1993). Building, Dwelling, Thinking. Basic Writings. London, Routledge. 344-363.

Henderson, R. S. (1990). "David Hume on Personal Identity and the Indirect Passions." Hume Studies **16**(1): 33-44.

Hertzberg, L. (1991). Imagination and the Sense of Identity. Human Beings. Cambridge, Cambridge University Press. 143-154.

Hirsch, E. (1991). "Divided Minds." The Philosophical Review **100**(1): 3-30.

Hollis, M. (1990). The Poetics of Personhood. Reading Rorty. Oxford, Basil Blackwell. 244-256.

Honderich Ted, e. (1995). The Oxford Companion to Philosophy. Oxford, Oxford University Press.

Howe, M. L. and M. L. Courage (1997). "The Emergence and Early Development of Autobiographical Memory." Psychological Review **104**(3): 499-523.

http://www.artistsuk.net/acatalog/ARTISTS_UK__PRINTS_133.html

Hume, D. (1888). Of Personal Identity. A Treatise of Human Nature. Oxford, Clarendon Press. 251-263.

Hume, D. (1888). A Treatise of Human Nature. Oxford, Clarendon Press.

Hume, D. (1962). Appendix (1740). A Treatise of Human Nature Book One. London, Collins. 324-331.

Hundert, E. J. (1992). "Augustine and the Sources of the Divided Self." Political Theory **20**(1): 86-104.

Husserl, E. (1960). Phenomenology. Realism and the Background of Phenomenology. Illinois, The Free Press of Glencoe. 118-128.

Ingold, T. (1995). Building, Dwelling, Living. Shifting Concepts. London, Routledge. 57-80.

James, W. (1952). The Consciousness of Self. The Principles of Psychology. Chicago, William Benton. 188-259.

Kant, I. (1929). Critique of Pure Reason (1787). London, MacMillan Press.

Keller, P. (1994). "Personal Identity and Kant's Third Person Perspective." Ideal Studies **24**(2): 123-146.

Kennett, J. and S. Matthews (2001). Identity, Control and Responsibility: the case of Multiple Personality Disorder. Time and Ethics, University of Otago, New Zealand.

Kenny, A. (1995). Frege. Harmondsworth, Penguin Books.

Kitcher, P. (1990). Kant's Transcendental Psychology. New York, N Y, Oxford University Press.

Kojeve, A. (1969). Introduction to the Reading of Hegel. Ithaca, N Y, Cornell University Press.

Kolak, D. (1993). "The Metaphysics and Metapsychology of Personal Identity." American Philosophical Quarterly **30**(1): 39-49.

Korsgaard, C. (1989). "Personal Identity and the Unity of Agency: A Kantian Response to Parfit." Philosophy & Public Affairs **18**(2): 101-132.

Krebs, W. A., Ed. (1981). Collins Australian Pocket Dictionary of the English Language. Sydney, Collins.

Kripke, S. A. (1980). Naming and Necessity. Oxford, Basil Blackwell.

Lacey, A. R. (1996). A Dictionary of Philosophy. London, Routledge.

Lampinen, J. M., J. S. Neuschatz, et al. (1998). "Memory Illusions and Consciousness: Examining the Phenomenology of True and False Memories." Current Psychology **16**(3/4): 181-224.

Lawler, R., D. W. Wuerl, et al., Ed. (1976). The Teaching of Christ. Huntington, Indiana, Our Sunday Visitor Inc.

Leiber, J. (1985). Can Animals and Machines be Persons? Indianapolis, Hackett Publishing Co.

Leibniz, G. W. (1973). Philosophical Writings. London, Everyman's Library.

Lewis, D. (1976). Survival and Identity. The Identities of Persons. Berkeley, CA, University of California Press. 17-40.

Lewis, D. (1997). Vague Identity: Evans Misunderstood. Vagueness: A Reader. Cambridge, MA, M I T Press. 318-320.

Little, William. (1933). The Shorter Oxford English Dictionary. Oxford, Clarendon Press.

Lloyd, G. (1993). Being in Time. London, Routledge.

Locke, D. (1971). Memory. London, Macmillan.

Locke, J. (1959). An Essay Concerning Human Understanding (1894). New York, N Y, Dover Publications Inc.

Locke, J. (1959). Of Identity and Diversity. An Essay Concerning Human Understanding (1894). New York, N Y, Dover Publications Inc. 439-470.

Loeb, L. E. (1992). "Causation, Extrinsic Relations, and Hume's Second Thoughts about Personal Identity." Hume Studies **18**(2): 219-231.

Lowe, E. J. (1991). Real Selves: Persons as a Substantial Kind. Human Beings. Ed Cockburn, David, Cambridge, Cambridge University Press. 87-107.

Lowe, E. J. (1991). "Substance and Selfhood." Philosophy **66**: 81-99.

Lowe, E. J. (1993). "Self, Reference and Self-Reference." Philosophy **68**: 15-33.

Luria, A. R. (1968). The Mind of a Mnemonist. New York, Basic Books.

Lyon, A. J. (1988). Problems of Personal Identity. An Encyclopedia of Philosophy. London, Routledge. 441-462.

Mach, E. (1959). The Analysis of Sensations. New York, N Y, Dover Publications, Inc.

MacIntyre, A. (1981). After Virtue. London, Duckworth.

MacIntyre, A. (1981). The Virtues, the Unity of a Human Life and the Concepts of a Tradition. London, Duckworth.

Macintyre, A. (1994). "Critical Remarks on 'Sources of the Self' by Charles Taylor." Philosophy and Phenomenological Research **54**(1): 187-189.

Mackie, J. L. (1976). Problems From Locke. Oxford, Clarendon Press.

Maddell, G. (1981). The Identity of the Self. Edinburgh, Edinburgh University Press.

Maddell, G. (1991). Personal Identity and the Idea of a Human Being. Human Beings. Cambridge, Cambridge University Press. 127-142.

Madell, G. (1985). "Derek Parfit and Greta Garbo." Analysis **45**: 105-109.

Malcolm, N. (1971). Problems of the Mind. London, Allen & Unwin.
Malcolm, N. (1977). Memory and Mind. London, Cornell University Press.

Malpas, J. (1999). "Constituting the Mind: Kant, Davidson and the Unity of Consciousness." International Journal of Philosophical Studies **7**(1): 1-30.

Malpas, J. (1999). Place and Experience. Cambridge, Cambridge University Press.

Margolis, J. (1988). "Minds, Selves, and Persons." Topoi (7): 31-45.

Margolis, J. (1995). "A Biopsy of Recent Analytical Philosophy." The Philosophical Forum **26**(3): 1-28.

Masson, J. and S. McCarthy (1996). When Elephants Weep. London, Vintage.

Matthews, G. B. (1992). Thought's Ego in Augustine and Descartes. Ithaca, N Y, Cornell University Press.

McDowell, J. (1994). Mind and World. Cambridge, MA, Harvard University Press.

McGeer, V. (1996). "Is "Self-Knowledge" an Empirical Problem? Renegotiating the Space of Philosophical Explanation." The Journal of Philosophy **93**(10): 483-515.

McInerney, P. K. (1991). "The Nature of a Person-Stage." American Philosophical Quarterly **28**(3): 227-235.

McIntyre, J. L. (1989). "Personal Identity and the Passions." Journal of the History of Philosophy **27**(4): 545-557.

McTaggart, J. (1927). The Nature of Existence. Cambridge, University Press.

Meinong, A. (1960). The Theory of Objects. Realism and the Background of Phenomenology. Illinois, The Free Press of Glencoe. 76-117.

Melin, R. (1998). Persons - Their Identity and Individuation. Umea, Sweden, Umea University.

Merleau-Ponty, M. (1962). Phenomenology of Perception. London, Routledge & Kegan Paul.

Merleau-Ponty, M. (1964). The Child's Relation with Others. The Primacy of Perception. Evanston, Northwestern University Press. 96-155.

Merleau-Ponty, M. (1964). The Primacy of Perception. Northwestern University Press.

Michel, G. F. and C. L. Moore (1995). Developmental Psychobiology. Cambridge, MA, MIT Press.

Millikan, R. (1993). Biosemantics. White Queen Psychology and other Essays. Cambridge, MA, MIT Press. 83-101.

Millikan, R. (1993). Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics. White Queen Psychology and Other Essays. Cambridge, MA, MIT Press. 123-133.

Millikan, R. (1993). Introduction. White Queen Psychology and other Essays. Cambridge, MA, MIT Press. 3-12.

Millikan, R. (1993). On Mentalese Orthography, Part 1. White Queen Psychology and other Essays. Cambridge, MA, MIT Press. 103-121.

Mills, E. (1993). “Dividing without Reducing: Bodily Fission and Personal Identity.” Mind **102**(405): 37-51.

Mitchell, H. (1991). “Charles Taylor on the Self, its Languages and its History.” History of Political Thought **12**(2): 335-358.

Nagel, T. (1970). Solipsism, Dissociation, and the Impersonal Standpoint. The Possibility of Altruism. Princeton, N J, Princeton University Press. 99-115.

Nagel, T. (1979). Brain Bisection and the Unity of Consciousness. Mortal Questions. Cambridge, Cambridge University Press. 147-164.

Nagel, T. (1979). Panpsychism. Mortal Questions. Cambridge, Cambridge University Press. 181-195.

Nagel, T. (1979). Subjective and Objective. Mortal Questions. Cambridge, Cambridge University Press. 196-213.

Nagel, T. (1979). What is it like to be a bat? Mortal Questions. Cambridge, Cambridge University Press. 165-180.

Nagel, T. (1986). The View from Nowhere. Oxford, Oxford University Press.

Neisser, U. (1982). John Dean's Memory: A Case Study. Memory Observed. San Francisco, W H Freeman and Company. 139-159.

Neisser, U. (1982). Memory Observed. San Francisco, W H Freeman and Company.

Neisser, U. (1986). Nested Structure in Autobiographical Memory. Autobiographical Memory. Cambridge, Cambridge University Press. 71-81.

Neisser, U. (1988). “Five Kinds of Self-Knowledge.” Philosophical Psychology **1**(1): 34-59.

Neisser, U. (1988). What is Ordinary Memory the Memory of? Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory. Cambridge, Cambridge University Press. 356-373.

Neisser, U. and E. Winograd, Ed. (1988). Remembering Reconsidered: Ecological and Traditional approaches to the study of Memory. Cambridge, Cambridge University Press.

Neuhouser, F. (1990). Fichte's Theory of Subjectivity. Cambridge, Cambridge University Press.

Nietzsche, F. (1967). The Birth of Tragedy and The Case of Wagner (1886). New York, Random House.

Noonan, H. (1993). “Chisholm, Persons and Identity.” Philosophical Studies **69**(1): 35-57.

Noonan, H. W. (1989). Personal Identity. London, Routledge.

Noonan, H., Ed. (1993). Personal Identity. Aldershot, Dartmouth Publishing Co.

Norton, P. B. and J. J. Esposito (1994). The New Encyclopaedia Britannica. Chicago, Encyclopaedia Britannica, Inc.

Nozick, R. (1981). The Identity of the Self. Philosophical Explanations. Oxford, Clarendon Press. 27-114.

Nozick, R. (1981). Philosophical Explanations. Oxford, Clarendon Press.

Oderberg, D. (1993). Fission, Intermittence and the Primitiveness of Identity. The Metaphysics of Identity over Time. New York, N Y, St. Martin's Press. 175-199.

O'Donohue, J. (1997). Anam Cara. Sydney, Bantam Books.

Oladipo, O. (1992). “The Yoruba Conception of a Person:
An Analytico-Philosophical Study.” International Studies in Philosophy **24**(3): 15-23.

Olafson, F. A. (1994). “Comments on 'Sources of the Self' by Charles Taylor.” Philosophy and Phenomenological Research **54**(1): 191-196.

Olson, E. (1992). “Was I Ever a Fetus?” Philosophy and Phenomenological Research **57**(1): 95-110.

Olson, E. (1994). “Is Psychology Relevant to Personal Identity?” Australasian Journal of Philosophy **72**(2): 173-186.

Olson, E. (1997). The Human Animal. Oxford, Oxford University Press.

Parfit, D. (1976). Lewis, Perry, and What Matters. The Identities of Persons. Berkeley, CA, University of California Press. 91-107.

Parfit, D. (1984). Appendices D Nagel's Brain. Reasons and Persons. Oxford, Clarendon Press. 468-477.

Parfit, D. (1984). Appendices E The Closest Continuer Schema. Reasons and Persons. Oxford, Clarendon Press. 477-479.

Parfit, D. (1984). How We Are Not What We Believe. Reasons and Persons. Oxford, Clarendon Press. 219-243.

Parfit, D. (1984). Personal Identity and Rationality. Reasons and Persons. Oxford, Clarendon Press. 305-320.

Parfit, D. (1984). Reasons and Persons. Oxford, Clarendon Press.

Parfit, D. (1984). What Does Matter. Reasons and Persons. Oxford, Clarendon Press. 281-306.

Parfit, D. (1984). What We Believe Ourselves To Be. Reasons and Persons. Oxford, Clarendon Press. 199-217.

Parfit, D. (1984). Why Our Identity Is Not What Matters. Reasons and Persons. Oxford, Clarendon Press. 245-280.

Parfit, D. (1992). “The Indeterminacy of Identity: A Reply to Brueckner.” :

Parkes, G. (1994). Composing the Soul. Chicago, University of Chicago Press.

Passmore, J. (1980). Hume's Intentions. London, Duckworth.

Peacocke, C. (1983). Sense and Content. Oxford, Clarendon Press.

Peacocke, C. (1992). A Study of Concepts. Cambridge, The MIT Press.

Pears, D. (1990). Hume's System. Oxford, Oxford University Press.

Pears, D. (1993). “Hume on Personal Identity.” Hume Studies **19**(2): 289-299.

Penelhum, T. (1976). Self-Identity and Self-Regard. The Identities of Persons. Berkeley, CA, University of California Press. 253-280.

Penelhum, T. (1992). “The Self of Book 1 and the Selves of Book 2.” Hume Studies **18**(2): 281-291.

Perry, J., Ed. (1975). Personal Identity. Topics in Philosophy. Berkeley, CA, University of California Press.

Perry, J. (1976). The Importance of being Identical. The Identities of Persons. Berkeley, CA, University of California Press. 67-90.

Perry, J. (1990). “Self-Notions.” Logos **11**: 17-31.
Perry, J. (1993). Castaneda on He and I. The Problem of the Essential Indexical and Other Essays. New York, Cambridge University Press. 91-119.

Perry, J. (1993). Frege on Demonstratives. The Problem of the Essential Indexical and Other Essays. New York, Oxford University Press. 3-32.

Perry, J. (1993). The Problem of the Essential Indexical. The Problem of the Essential Indexical and Other Essays. New York, Oxford University Press. 33-52.

Perry, P. (1999). Making Strides in the Battle Against Alzheimer's Disease. The Saturday Evening Post. Indianapolis, IN, 46-47; 65-69.

Petrik, J. (1989). “Two Faces Have 'I'.” Grazer Philosophical Studies **36**: 55-65.

Piper, A. M. (1985). “Two Conceptions of the Self.” Philosophical Studies **48**: 173-198.

Plato (1959). Wisdom and the Soul: Phaedo. The Last Days of Socrates. Harmondsworth, Penguin. 93-191.

Plato (1974). The Republic. Harmondsworth, Penguin Books.

Popplestone, J. A. and M. W. McPherson (1988). Dictionary of Concepts in General Psychology. New York, Greenwood Press.

Potter, N. (1996). “Loopholes, Gaps, and What is Held Fast.” Philosophy, Psychology, and Psychiatry **3**(4): 237-254.

Proust, M. (1981). Remembrance of Things Past. Harmondsworth, Penguin.

Puccetti, R. (1968). Persons. London, Macmillan.

Putnam, H. (1981). Reason, Truth, and History. Cambridge, Cambridge University Press.

Quine, W. V. (1969). Ontological Relativity. Ontological Relativity and Other Essays. New York, Columbia University Press. 36-68.

Quine, W. V. (1996). Two Dogmas of Empiricism (1951). The Philosophy of Language. Oxford N Y, Oxford University Press. 3rd., ed. 30-52.

Quine, W. V. O. (1960). Word and Object. Cambridge, MA, MIT Press.

Radice, B. (1961). The Cloud of Unknowing and Other Works. London, Penguin.

Rawls, J. (1971). A Theory of Justice. Cambridge, MA, The Belknap Press.

Real, R. (1995). “Foucault contra Taylor: Whose Sources? Which Self?” Dialogue **34**(4): 663-674.

Reber, A. S. (1985). The Penguin Dictionary of Philosophy. Harmondsworth, Viking.

Reed, E. S. (1996). Encountering the World. Oxford, Oxford University Press.

Reid, T. (1785). Essays on the Intellectual Powers of Man. Edinburgh, Scolar Press.

Reid, T. (1803). An Enquiry into the Human Mind. Account of the Life and Writings of Thomas Reid. Edinburgh, Maclachlan & Stewart.

Ricoeur, P. (1991). "Narrative Identity." Philosophy Today **35**(1): 73-81.

Ricoeur, P. (1992). Oneself as Another. Chicago, University of Chicago Press.

Robinson, J. (1986). Autobiographical Memory: A Historical Prologue. Autobiographical Memory. Cambridge, Cambridge University Press. 19-24.

Robinson, J. (1988). "Personal Identity and Survival." The Journal of Philosophy **85**(6): 319-328.

Rolle, R. (1972). The Fire of Love. London, Penguin.

Rorty, A. O. (1976). Introduction. The Identities of Persons. Berkeley, University of California Press. 1-15.

Rorty, A. O. (1976). A Literary Postscript: Characters, Persons, Selves, Individuals. The Identities of Persons. Berkeley, CA, University of California Press. 301-323.

Rosenberg, J. F. (1988). "On Not Knowing What or Who One Is: Reflections on the Intelligibility of Dualism." Topoi **7**: 57-63.

Rosenfield, I. (1992). The Strange, Familiar and Forgotten. London, Picador.

Rosenow, E. (1989). "Kierkegaard's Existing Individual." Journal of Philosophy of Education **23**(1): 3-13.

Rosner, M. I. and M. E. Raichle (1994). Images of Mind. New York, Scientific American Library.

Rovane, C. (1987). "The Epistemology of First-Person Reference." The Journal of Philosophy **84**(3): 147-167.

Rovane, C. (1998). The Bounds of Agency. Princeton, NJ, Princeton University Press.

Rubin, D. C. (1986). Autobiographical Memory. Cambridge, Cambridge University Press.

Russell, J. (1995). At Two with Nature: Agency and the Development of Self-World Dualism. The Body and the Self. Cambridge, M A, MIT Press. 127-151.

Sacks, O. (1985). The Man who Mistook his Wife for a Hat. London, Pan Books.

Sainsbury, R. M. (1995). Paradoxes. Cambridge, Cambridge University Press.

Schacter, D. L., M. Verfaellie, et al. (1998). "When True Recognition Suppresses False Recognition: Evidence from Amnesic Patients." Journal of Cognitive Neuroscience **10**(6): 668-679.

Schechtman, M. (1990). "Personhood and Personal Identity." The Journal of Philosophy **87**(2): 71-92.

Schechtman, M. (1994). "The Truth about memory." Philosophical Psychology **7**(1): 3-18.

Schechtman, M. (1997). "The Brain/Body Problem." Philosophical Psychology **10**(2): 149-164.

Scheffler, I. (1967). Science and Subjectivity. Indianapolis, Bobbs-Merrill Co Inc.

Schmidt, R. F., Ed. (1978). Fundamental of Neurophysiology. New York, Springer-Verlag.

Schutte, O. (1991). "Irigaray on the Problem of Subjectivity." Hypatiia **6**(2): 64-76.

Sculthorpe, P. (1999). Sun Music. Sydney, ABC Books.

Searle, J. R. (1997). The Mystery of Consciousness. London, Granta Books.

Seitz, B. (1991). "The Identity of the Subject, After Sartre." Philosophy Today **35**(4): 362-370.

Sellars, W. (1997). Empiricism and the Philosophy of Mind. Cambridge, MA, Harvard University Press.

Seth, A. (1976). The Development from Kant to Hegel. New York, Garland Publishing Inc.

Sheets-Johnstone, M. (1999). On Learning to Move Oneself. The Primacy of Movement. Amsterdam, John Benjamins. 223-272.

Shoemaker, S. (1963). Self-Knowledge and Self Identity. Cornell University Press.

Shoemaker, S. (1976). Embodiment and Behavior. The Identities of Persons. Berkeley, CA, University of California Press. 109-137.

Shoemaker, S. (1984). Personal Identity: A Materialist's Account. Personal Identity. Oxford, Basil Blackwell. 67-132.

Shoemaker, S. (1984). Persons and their Pasts. Identity, Cause, and Mind. Cambridge, Cambridge University Press. 19-48.

Shoemaker, S. (1984). Self-reference and Self-awareness. Identity, Cause, and Mind. Cambridge, Cambridge University Press. 6-18.

Shoemaker, S. (1992). "Unger's Psychological Continuity Theory." Philosophy and Phenomenological Research **52**(1): 139-143.

Shoemaker, S. S. (1969). Comments. Perception and Personal Identity. Cleveland, OH, The Press of Case Western Reserve University. 107-127.

Sinclair, A. S. (1995). "The Failure of Thomas Reid's Attack on David Hume." British Journal of the History of Philosophy **3**(2): 389-398.

Skinner, Q. (1991). "Who are 'We'? Ambiguities of the Modern Self." Inquiry **34**(2): 133-153.

Slater, A. (1989). Visual Memory and Perception in Early Infancy. Infant Development. Hove, UK, Lawrence Erlbaum Associates. 43-72.

Slater, A. and G. Bremner, Ed. (1989). Infant Development. Hove, U K, Lawrence Erlbaum Associates.

Slors, M. (1998). "Two Conceptions of Psychological Continuity." Philosophical Explorations (1): 61-80.

Smith, A. D. (1998). Ageing of the Brain: Is Mental Decline Inevitable? From Brains to Consciousness. London, Penguin. 94-108.

Smook, R. (1986). "Does Remembering doing the deed presuppose Personal Identity?" Dialogue **25**(2): 363-366.

Snowdon, P. F. (1991). Personal Identity and Brain Transplants. Human Beings. Cambridge, Cambridge University Press. 109-126.

Soll, I. (1969). An Introduction to Hegel's Metaphysics. Chicago, University of Chicago Press.

Soloman, P. R., G. R. Goethals, et al., Ed. (1989). Memory: Interdisciplinary Approaches. New York, Springer-Verlag.

Sprigge, T. L. S. (1988). "Personal and Impersonal Identity." Mind **97**: 29-49.

Stephens, G. L. and G. Graham (1994). "Self-Consciousness, Mental Agency, and the Clinical Pathology of Thought Insertion." Philosophy of Psychiatric Psychology **1**(1): 1-10.

Stern, C. (1987). "Hume and the Self at a Moment." History of Philosophy Quarterly **4**(2): 217-233.

Stern, D. (1985). The Interpersonal World of the Infant. New York, N Y, Basic Books.

Stich, S. P. (1983). From Folk Psychology to Cognitive Science. Cambridge, MA, MIT Press.

Storl, H. (1992). "The Problematic Nature of Parfitian Persons." Personalist Forum **8**(1): 123-132.

Strawson, G. (1989). Hume on the Mind. The Secret Connexion. Oxford, Clarendon Press. 129-134.

Strawson, G. (1996). “The Sense of Self.” London Review of Books (April, 1996): 21-23.

Strawson, G. (1997). “'The Self'.” Journal of Consciousness Studies **4**(5-6): 405-428.

Strawson, P. F. (1959). Individuals. London, Methuen.

Strawson, P. F. (1966). The Bounds of Sense. London, Methuen.

Strawson, P. F. (1992). “Comments on Some Aspects of Peter Unger's 'Identity, Consciousness and Value'.” Philosophy and Phenomenological Research **52**(1): 145-147.

Stroud, B. (1977). Hume. London, Routledge & Kegan Paul.

Stumpf, S. E. (1994). Philosophy and its Problems. New York, McGraw-Hill Inc.

Swain, C. (1991). “Being Sure of One's Self: Hume on Personal Identity.” Hume Studies **17**(2): 107-124.

Swinburne, R. (1984). Personal Identity: the Dualist Theory. Personal Identity. Oxford, Basil Blackwell. 1-66.

Swinburne, R. (1992). “Discussion of Peter Unger's 'Identity, Consciousness and Value'.” Philosophy and Phenomenological Research **52**(1): 149-153.

Swinburne, R. (1997). The Evolution of the Soul. Oxford, Clarendon Press.

Tanner, M. (1994). Nietzsche. Oxford, Oxford University Press.

Taylor, C. (1994). “Precis of 'Sources of the Self'.” Philosophy and Phenomenological Research **54**(1): 185-186.

Taylor, C. (1994). “Reply to Commentators.” Philosophy and Phenomenological Research **54**(1): 203-213.

Thornton, M. (1991). “Same Human Being, Same Person?” Philosophy **66**(January): 115-118.

Tortora, G. J. (1980). Principles of Human Anatomy. New York, Harper & Row.

Tortora, G. J. and J. F. Becker (1978). Life Science. New York, MacMillan Publishing Co.

Trendelenburg, A. (1910). “A Contribution to the History of the Word Person.” Monist **20**: 336-363.

Trigg, R. (1988). “The Metaphysical Self.” Religious Studies **24**: 277-289.

Turner, G. W. (1978). The Australian Pocket Oxford Dictionary. Melbourne, Oxford University Press.

Unger, P. (1988). Conscious Beings in a Gradual World. Midwest Studies in Philosophy, X11. Minneapolis, University of Minnesota Press. 287-333.

Unger, P. (1990). Identity, Consciousness and Value. New York, Oxford University Press.

Unger, P. (1992). “Precis of 'Identity, Consciousness and Value'.” Philosophy and Phenomenological Research **52**(1): 133-137.

Unger, P. (1992). “Reply to Reviewers.” Philosophy and Phenomenological Research **52**(1): 159-177.

Unger, P. (1994). “Critical Notice.” Canadian Journal of Philosophy **24**(1): 119-134.

Uzgalis, W. L. (1990). “Relative Identity and Locke's Principle of Individuation.” History of Philosophy Quarterly **7**(Number 3): 283-297.

Van Gelder, T. (1995). “What Might Cognition be, if not Computation?” The Journal of Philosophy **91**(7): 345-381.

Velleman, J. D. (1996). “Self to Self.” The Philosophical Review **105**(1): 39-77.

Vesey, G. (1977). Personal Identity. Ithaca, NY, Cornell Paperbacks.

Vesey, G. N. (1965). The Embodied Mind. London, Allen & Unwin.

Vico, G. (1982). The Idea of a Dictionary of Mental Words which is Common to All Nations. Selected Writings. Cambridge, Cambridge University Press. 151-153.

Vigotsky, L. S. (1962). Thought and Language. Cambridge, N A, M I T Press.

Von Uexkull, J. (1934). A Stroll Through the Worlds of Animals and Men. Instinctive Behavior. New York, International Universities Press, Inc. 5-128.

Waddington, C. H. (1966). Principles of Development and Differentiation. New York, The Macmillan Company.

Walker, R. C. S. (1978). Kant. London, Routledge & Kegan Paul.

Ward, A. (1985). “Hume, Demonstratives, and Self-Ascriptions of Identity.” Hume Studies **11**(April): 69-93.

Warren, M. A. (1994). On the Moral and Legal Status of Abortion. Contemporary Issues in Bioethetics. Belmont, CA, Wadsworth Publishing. 4th, ed. 302-311.

Waxman, W. (1992). “Hume's Quandary Concerning Personal Identity.” Hume Studies **18**(2): 233-253.

Wedeking, G. (1987). “Locke's Metaphysics of Personal Identity.” History of Philosophy Quarterly **4**(Number 1): 17-31.

Wedeking, G. (1990). “Locke on Personal Identity and the Trinity Controversy of the 1690s.” Dialogue **29**(Number 2): 163-188.

Weiten, W. (1992). Human Memory. Psychology, Themes and Variations. Pacific Grove, California, Brooks/Cole. 2nd, ed. 230-267.

West, M. (1996). A View from the Ridge. Sydney, Harper Collins.

Wiggins, D. (1967). Identity and Spatio-Temporal Continuity. Oxford, Basil Blackwell.

Wiggins, D. (1980). Sameness and Substance. Oxford, Basil Blackwell.

Wilkes, G. A. and W. A. Krebs (1995). Collins Concise Dictionary. Sydney, Harper Collins.

Wilkes, K. (1988). Real People. Oxford, Clarendon Press.

Wilkes, K. V. (1991). How Many Selves Make Me? Human Beings. Ed Cockburn, David, Cambridge, Cambridge University Press. 235-243.

Will, F. L. (1968-69). “Thoughts and Things.” Proceedings and Addresses of the American Philosophical Association **42**: 51-69.

Williams, B. (1973). Are Persons Bodies? Problems of the Self. Cambridge, Cambridge University Press. 64-81.

Williams, B. (1973). Bodily Continuity and Personal Identity. Problems of the Self. Cambridge, Cambridge University Press. 19-25.

Williams, B. (1973). Imagination and the Self. Problems of the Self. Cambridge, Cambridge University Press. 26-45.

Williams, B. (1973). Personal Identity and Individuation. Problems of the Self. Cambridge, Cambridge University Press. 1-18.

Williams, B. (1973). Problems of the Self. Cambridge, Cambridge University Press.

Williams, B. (1973). The Self and the Future. Problems of the Self. Cambridge, Cambridge University Press. 46-63.

Williamson, T. (1994). Vagueness. London, Routledge.

Winkler, K. P. (1991). "Locke on Personal Identity." Journal of the History of Philosophy **29**(Number 2): 201-226.

Wittgenstein, L. (1958). The Blue and Brown Books. Oxford, Basil Blackwell.

Wittgenstein, L. (1958). Philosophical Investigations. Oxford, Basil Blackwell.

Wittgenstein, L. (1961). Tractatus Logico-Philosophicus. London, Routledge & Kegan Paul.

Wojtyla, K. (1969). The Acting Person. Dordrecht, D. Reidel Publishing Company.

Wright, J. P. (1983). The Sceptical Realism of David Hume. Minneapolis, University of Minnesota Press.

Yandell, K. E. (1992). "Continuity, Consciousness, and Identity in Hume's Philosophy." Hume Studies **18**(2): 255-274.

Yolton, J. (1993). A Locke Dictionary. Oxford, Blackwell.

Zoeller, G. (1989). "Making Sense Out of Inner Sense: The Kantian Doctrine as Illuminated by the Leningrad 'Reflexion'." International Philosophical Quarterly **29**(3): 263-270.