University of Alberta

Personal Identity as a Social Concept

by

Edward Hayman

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Philosophy

©Edward Hayman
Fall 2010
Edmonton, Alberta

Examining Committee

Wesley Cooper, Philosophy

Bernard Linsky, Philosophy

Adam Morton, Philosophy

Leo Mos, Psychology

DEDICATION

This work is dedicated to my wife, Stephanie, whose support throughout the course has been indispensible, and to my daughters, Jessica and Héloïse. They have all been loyal, interested, positive and inspiring.

ABSTRACT

The Thesis argues that the concept of 'personal identity' is developed in social circumstances, relating to ideas about how the self continues through time and to 'person-directed' concerns. Chapter one uses William James's classification of the constituents of the self, and his idea of the 'stream of consciousness', as an introduction to the concept. Chapters two and three deal with: George Mead's ideas about the self arising in social interaction; Eric Olson's distinction between 'biological' and 'psychological' continuity; Mark Johnston's view of 'wide psychological continuity' and his 'relativist' approach to personal identity; Robert Nozick's notions of 'reflective self-reference' and the 'closest continuer'; Derek Parfit's 'reductionist' approach; Wesley Cooper's elaboration of Nozick's account. Chapter four favours a physicalist account of the self and a flexible approach to the concept of personal identity, accommodating the needs and practices of the society in which the individual finds himself.

PREFACE

Thesis

'Personal identity' is a 'social concept' in that how an individual's identity is formed and sustained depends, not just on that individual, but on the people around him and interaction with them. Emphasis on the social aspects of the concept will help to discriminate among current philosophical approaches to the topic and to suggest ways in which theories of personal identity may be developed.

The thesis is motivated by Robert Nozick's account of personal identity in *Philosophical explanations* (1981). According to my 'hard' interpretation, Nozick claims that an individual achieves personal identity by his own isolated act and chooses what constitutes continuation for his identity. Nozick's account is analyzed in chapter three. It runs counter to the notion that personal identity is developed in social circumstances, and ignores the extent to which the concept is essential for the society surrounding an individual as well as for the individual himself. In Kafka's short story, *The Metamorphosis*, Gregor Samsa wakes up to find himself transformed into a 'verminous bug' (1915/2009). The tale is used by Wesley Cooper in his 2008 paper, *An Eldritch Tale: Kafka's 'The Metamorphosis' and the Self* (2008a, also discussed in chapter three); it illustrates social problems connected with personal identity, as well as individual ones. While Cooper is concerned with how Gregor might or might not accept his new embodiment, much of the story is about the affect of the metamorphosis on Gregor's family. William James in *The principles of psychology* (1890), analyzed in chapter one, shows the importance of social 'roles' to the formation of the 'empirical' self. We learn to behave differently in different social circumstances and the disposition to do this becomes part of our personal identity. A thought experiment in my paper suggests that an infant surviving in an uninhabited world would not develop personal identity, at least as we understand it. So far as I know this story has yet to be written as a piece of fiction. Writers have, of course,

speculated about how people would develop deprived of human society. Romulus and Remus, the fabled founders of Rome brought up by a wolf, supposedly survived quite well, as did Kipling's Mowgli, also brought up by animals. Real feral children seem to fare rather worse than fictional ones (see, for example, http://www.feralchildren.com). The philosophical claim I want to emphasize, made by George Mead in *Mind, Self, and Society* (1934, p 135), is that the 'self' 'arises in the process of social experience and activity'. My paper will suggest that Nozick's account is open to a 'softer' interpretation, allowing the recognition of social elements. Cooper, who is influenced by Nozick, provides a bridge between the 'social' and 'individual' approaches to personal identity. I argue that recognizing the social elements of the concept is essential for understanding how it is used. Social interaction is a necessary condition for becoming a person in the first place and for normal personal development.

ACKNOWLEDGEMENT

Table of Contents

List of Tables

INTRODUCTION

The defense of this thesis requires me to travel across territory much trampled upon and muddied. I use a supporting structure, which consists partly in selecting philosophers upon whose shoulder to lean, partly in trying to simplify the concepts with which I am dealing.

Concepts

I will mention some of the concepts here, as a guide to my use of particular terms:

I interpret 'society' in the widest possible way to mean any group of people reacting to an individual, extending from family through communities and nations to the world-population (that has a stake in the development of global figures). As children we develop into mature persons in the context of our immediate family. But, depending on the individual, no social layer is too distant or too great to lack all potentiality for influence. Mead (1934, p 157) declares that the individual's relations with social communities of diverse functions include 'individual members from several (in some cases from all) such communities.'

I will refer to the person whose personal identity we are considering as the 'individual'. Society's capacity to influence personal identity varies according to the characteristics of the society, the characteristics of the individual and the latter's stage of development. William James, for example, in a passage (1890, p 315) mentioned in chapter one, compares the youth who courts social recognition with the mature man who may defy social mores in pursuit of his own ideals.

The concept of 'personal identity' is associated with that of 'self'. What matters for the discussion of personal identity is:

- 'continuity' - how the 'self' continues through time. What makes the individual the same person now as previously or as in the future?
- 'care' – what Tamar Gendler (2009) calls the 'practical and emotional significance' of being the same person through time, significance both for the individual and society.

The concept of 'self' defies simplification. A difficulty is that the word is used in so many different contexts to refer to both physical and non-physical phenomena. Indeed, it can be argued that there is no 'concept' of self in the singular and that the word is just a linguistic convenience, often used as a suffix for reflexive reference, as in 'Know thyself'. This command could be translated roughly as 'Know what sort of a person you are'. It would be possible to write about personal identity (clumsily) without using the word 'self'. However, the philosophers to whom I mainly refer assume that there is a concept of 'self' relevant to the concept of 'personal identity' and I shall make the same assumption. One way of putting my position is to say that I treat the existence of the self as a 'logical postulate'. A 'logical postulate' is a proposition treated as not requiring proof, either because it is undeniable or because it is necessary as the basis for the logical deduction of other propositions. The term recurs frequently in this paper. I take logical postulates to be aids to understanding rather than statements of fact, which can be proved, or statements of non-evidential beliefs, which are justified differently. (Non-evidential beliefs are considered in chapter three's 'Background point on mysteries and beliefs'.) The distinction between 'logical postulates' and 'non-evidential beliefs' is, perhaps, not always clear. Paradigm cases aid clarity. The post-Kantian concept of the 'absolute I', which helps with the understanding of self-consciousness, is a paradigm case of a 'logical postulate'. The belief in a 'soul' created from a non-physical substance is a paradigm case of a non-evidential belief. The 'soul' is regarded by believers, like Kant, as a '*noumenon*', an unprovable 'fact' inaccessible to humans.

Problems with regard to the self are enhanced by the difference between the individual's view of himself, via self-consciousness, and society's external view of the individual. The metaphysical question, 'What is the self?', remains unanswered. I consider a range of approaches, with the Cartesian at one end and the physicalist at the other. For Descartes the self is essentially a 'thinking thing' – thinking being an attribute of the soul – and a separate substance[1] from that of the body (1641/1993, p 19). Galen Strawson describes the physicalist version of the self as whatever system there is 'in the brain that supports the consciousness and personality of a human being' (2009, p 562). In between, there are various approaches from those who are dissatisfied with a purely physical account of the self but fall short of positing an independent mental substance.

It is the relationship between personal identity and self that occupies the foreground of this paper. In the background are the ideas of 'personhood' and 'personality'. Only a human animal can be a 'person' and he has to be one to achieve personal identity. How and when a human being becomes a person is important for my discussion. 'Personality' is more of a common-sense concept than a philosophical one; I take it to comprise the outward characteristics of a person, accessible to society and not just to the individual concerned. (I shall not worry here about the question whether a human being, or even a non-human animal, can have 'personality' without being a person.)

Philosophers
Chapter one is a critical analysis of William James's account of personal identity, largely focused on chapter ten of *The Principles of Psychology* (1890). This gives me a framework within which to establish more firmly the concepts mentioned above, and enables me to introduce ideas about personal identity and the self which are important later on. As a platform for the main defense of my thesis, I set up a fairly artificial opposition. There are philosophers who

---

[1] This paper does not give a full explanation of 'substance' but a modern take on the concept is presented in the section on Johnston in chapter 2.

emphasize the role of society, in both forming the concept of personal identity and developing individual personal identities. Chapter two deals with George Mead, Eric Olson, and Mark Johnston as examples of 'social' theorists. Chapter three deals with Robert Nozick, Derek Parfit, and Wesley Cooper as examples of those who emphasize the role of the individual in establishing his own personal identity and deciding what constitutes continuity for him as a person. The contrast is not intended to suggest two irreconcilable warring factions among philosophers of mind. It simply serves the purpose of this paper. The contrast is also not intended to suggest that either the individual or society is totally responsible for the aspects of self conceived as personal identity. Much of what a person cares about and what constitutes his continuity is simply a fact of life, not inexplicable but equally not controllable. However, the contrasting and overlapping contributions of both the individual and society are of interest. Chapter four tries to clarify some of the muddier issues raised in the previous chapters and to gather the main threads of my defense. 'Conclusions' may seem too strong a word for the set of tentative suggestions grouped under this sub-heading. The personal identity concept remains 'loose', as suggested in the chapter two discussion of Johnston, awaiting further scientific investigation and further 'philosophical explanations'.

CHAPTER 1: ENTRANCE THROUGH WILLIAM JAMES

William James's account

To defend my thesis it is important to have a background picture of the concept of 'self' in order to see how elements of that concept contribute to personal identity. William James provides a way of establishing this picture. In *The Principles of Psychology*, published in 1890, he writes extensively about the self. His writing precedes the considerable discussion of personal identity during the twentieth century, and I can use it as a setting against which to examine the conflicting twentieth-century arguments. This chapter is a critical analysis of James, largely of Chapter X in *The Principles of Psychology*.

James's very broad account includes elements that some common-sense as well as philosophical interpretations would not count as part of the self. 'In its widest possible sense, however, a man's Self is the sum total of all that he CAN call his' (1890, p 291). James's suggestion is that material possessions, such as clothes and property, are not merely indicative of personal identity but are a part of the self. He would have understood the sentiment of the Mack David/Harold Grant song: 'If I can't take it with me when I go, I just ain't gonna go.' James, however, while regarding these material possessions as 'part' of the self, does not consider them 'essential' in the way that the 'spiritual' self is. The non-empirical aspect of the 'spiritual' self will receive more detailed treatment in this paper. Meanwhile, it can be said that James's broadness is an advantage for a wide-ranging discussion. Another advantage is that he adopts a classificatory approach, which provides an orderly method of unpacking a complex and amorphous concept.

The empirical self

I begin, as James does (1890), with a discussion of what he calls the 'empirical Self'. He tabulates his classification of this on p 329. (I paraphrase his category descriptions.):

Table 1. James's classification of the 'empirical self'

	'Material'	'Social'	'Spiritual'
'Self-seeking'	Bodily appetites and instincts. Love of material possessions, such as clothes and home	Love, sociability, ambition, envy, pursuit of honour	Intellectual, moral and religious aspirations.
'Self-estimation'	Vanity, pride, modesty based on possessions	Social and family pride, snobbery, humility, shame	Sense of mental or moral superiority or inferiority

James's interpretation of 'empirical' data, known by experience, includes the results of introspection as well as external perception. He includes, as can be seen from the table, long-term feelings and tendencies, such as vanity and pride. He writes: 'The Empirical Self of each of us is all that he is tempted to call by the name of *me*' (p 291).

He says that the empirical 'constituents' of the self are three:

1. The 'material'
2. The 'social'
3. The 'spiritual'.

1. The 'material'

This ranges from the person's body, 'the innermost part of the material Self in each of us', to his occupations. In between these boundaries of the self's material field come his clothes, his family and his home. 'Our immediate family is part of ourselves' (1890, p 292). His possessions and his achievements are included.

'The parts of our wealth most intimately ours are those which are saturated with our labour' (p 293). The destruction of a 'life-long construction' such as a collection of insects or a manuscript would cause 'a sense of the shrinkage of our personality'.

Taken literally, this description of the 'material' constituent of the self covers far too much. I am not tempted to call my achievements or my home, my clothes or even my family 'me'. On the contrary, I have from an early age to distinguish my self from the objects and other people that surround me. Of the panoply of items which James attributes to the 'material' constituent only the body, including the brain, corresponds to the everyday notion of the self. Yet, seen from the point of view of personal identity, these material elements of the self's environment do have a special significance. I do identify with my family and my home. I do particularly care about my achievements and even, to the extent that I would not wish to be seen walking around in rags, about my clothes. These material things are important to me because they reflect my personality. I do not buy clothes which I think do not suit me. I might disown a child whose behaviour sufficiently transgressed what I expect from a member of my family.

James here seems to have identified material things which are not part of the self but are associated with the self in such a way that a person suffers if they are damaged or experiences happiness if they prosper. Not constituents of personal identity, they are part of the context in which personal identity develops.

2. The 'social'

For James a person's social self is 'the recognition he gets from his mates' (1890, p 293). There could be 'no more fiendish punishment', he supposes, than being totally unnoticed in society. He goes on to suggest that a person has more than one social self, as many, in fact, as there are individuals or individual groups who 'recognize' him and for whose opinion he cares. James emphasizes how a person

may differ according to his social environment, treating, say, his employees one way and his family quite differently. Worrying about how he appears to one group and not how he appears to another. 'What may be called "club-opinion" is one of the very strongest forces in life' (p 295).

That the social self is important is one of the main tenets of this paper. An individual's concept of himself is influenced by how others see him and is constructed partly of their concepts of how he is and how he should be. Here James is concerned with how a person 'shows different sides of himself' to different groups (p 294). A youth who is 'demure' with his parents and teachers may swear and swagger among his friends. James is also concerned with situations in which different social circumstances elicit different behaviour. A doctor or a priest will stay on in a plague-ridden city when ordinary people flee. The Nozickean approach to personal identity, which I shall be discussing, suggests that the individual chooses who he will be. Certainly 'choice' is an apt word for describing the acquisition of many empirical aspects of personal identity. The individual chooses his clothes, his job, his friends. But social groups also exercise choice about which aspects of personal identity they will recognize and which persons they will accept. There is interaction and personal identity develops in a social context.

3. The 'spiritual'

James distinguishes between the 'empirical' spiritual self, a 'man's inner or subjective being, his psychic faculties or dispositions' (1890, p 296) and the non-empirical 'pure Ego', to be dealt with later. With the 'spiritual' self we have come to what James regards as essential to the self. If it is altered a man becomes '*alienatus a se'*. The implication is that the 'material' and 'social' constituents, including the body, could change without a man losing his personal identity. Not so with dispositions such as 'ability to argue and discriminate' and 'moral sensibility and conscience'. We derive a 'purer self-satisfaction' from these than from the other constituents of the self. Such a division of the 'spiritual' self into

faculties, 'isolating them one from another, and identifying ourselves with either in turn', James characterizes as an 'abstract' approach. In consciousness 'as it actually presents itself', he says, 'a plurality of such faculties are always to be simultaneously found'.

Notice that these dispositions , which are so crucial to the self, are accessible to other people as well as to the individual whose self is under consideration. The latter may have his own privileged access to the thoughts and feelings which are associated with his various tendencies, but others will certainly be aware of his 'ability to argue' and his 'moral sensibility'. If his words and actions consistently contradict his claim to, say, 'moral sensibility', either he is lying or he is deluded. While not denying this 'shared' aspect of the empirical self, James's emphasis is on the self-conscious aspect of it. He describes consideration of the self, 'whether we take it abstractly or concretely', as 'a reflective process', the result of our ability '*to think of ourselves as thinkers*' (p 296).

James suggests that, in addition to this 'abstract' approach to the 'spiritual' self, there is a 'concrete' approach through what he calls 'the stream of consciousness'. He introduces this concept of the 'stream of consciousness' or 'stream of thought' in Chapter IX of 'The Principles of Psychology'[2]. He says that while the objects of consciousness are separate, whether they are external or internal to the conscious being, consciousness itself is continuous. It 'flows' like a river or stream (p 239). In an individual's sequence of thoughts, the separate thoughts are not independent of each other. Each thought is identified by its content – James says we 'name' each thought after its 'thing' (p 241) – but the contents of numerous other thoughts are always in the background. 'Whatever the content of the ego may be, it is habitually felt *with* everything else by us humans and must

---

[2] Chapter IX in the 1890 edition of 'The Principles of Psychology' is entitled 'The Stream of Thought'. In the revised, much shorter, version of the book published as 'Psychology, Briefer Course' in 1892, the equivalent chapter is entitled 'The Stream of Consciousness'. According to the Oxford English Dictionary the phrase first appeared in a publication by A.Bain in 1855. James, to whom the origin is sometimes attributed (see eg Wikepedia, http://en.wikipedia.org/wiki/Stream_of_consciousnes), may have popularized the phrase.

form a *liaison* between all the things of which we become successively aware' (p 242).

To return to Chapter X, James develops his 'concrete' approach to the 'spiritual' self by identifying it with the 'entire stream of our personal consciousness or the present "segment" or "section" of that stream, according as we take a broader or narrower view' (1890, p 296). That James allows either view is important in a discussion of personal identity. Some contemporary philosophers emphasize the need for 'psychological continuity' (e.g. Parfit, 1984, and Sydney Shoemaker, 1984); others adopt what Galen Strawson (2009, p 551) calls the 'thin' conception of personal identity. In any case, for James both the stream and the section are 'concrete existences in time' (p 296). And these 'concrete existences' are more clearly identified with the self than the material and social constituents. Yet, according to James, a '*certain portion of the stream abstracted from the rest*' is particularly identified with the self. It is a sort of 'innermost centre within the circle', a 'sanctuary within the citadel' (p 296) which is the centre of conscious control. It is the 'source of effort and attention', 'a sort of junction at which sensory ideas terminate and from which motor ideas proceed' (p 298).

This 'self of all the other selves', as he calls it (p 297), is not graspable in the way that the stream from which it comes can be perceived. He claims that 'all men' would agree 'up to a certain point (p 297) with the way he has identified it and described it so far. But there would be, he acknowledges, disagreement about what it is. Some would identify it with the soul; others would identify it with the imaginary denotation of the pronoun 'I'; and there would be all sorts of opinions in between. In this sense the 'sanctuary' abstraction (abstracted from the rest of the stream of consciousness) differs from that of the dispositions. Anyone can observe the different dispositions in so far as they manifest themselves in behaviour. In some ways it is easier to observe the abstracted dispositions than the complex whole self from which they have been abstracted.

James asserts that calling it an 'abstraction' does not mean 'it could not be presented in a particular experience' but only means it is not 'found alone' in the stream of consciousness. In a virtuoso display of introspection, he goes on to describe how he experiences his 'central active self' (p 299). He describes the 'constant play of furtherances and hindrances in my thinking'. He describes the way the thought processes are felt as 'as movements of something in my head' (p 300); 'pressures, convergences, divergences, and accommodations in my eyeballs'; 'the opening and closing of the glottis' (p 301); and other physical reactions. Yet he is careful to present all this as how he personally feels, without trying to generalize from it, and to emphasize that all he can draw from it is a hypothesis.

What James hypothesizes is that the 'feeling of spiritual activity' is a 'feeling of bodily activities' (p 302). The 'nuclear' part of the self is then a collection of physiological acts, 'adjustments', the 'less intimate, more shifting self' being executions. The adjustments 'few in number, incessantly repeated, constant amid great fluctuations in the rest of the mind's content' do not attract detailed introspective attention. Yet they are 'a coherent group of processes' contrasting with everything else in consciousness. They are 'primary reactions' acting as the 'permanent core of turnings-towards and turnings-from ' (p 303). If these 'primary reactions' constitute the 'ultimate' (p 304) self among the selves we can experience, it follows that all that we experience is 'objective'. The objective divides into self and not-self and there is nothing else, except the fact that these are known (by experience) and the stream of consciousness exists as the 'indispensable subjective condition' of this experience. James then points out that the stream of consciousness is not itself known immediately. 'It is only known in subsequent reflection' (p 304). It is not '"thinking its own existence along with whatever else it thinks" (as Ferrier says)'. James suggests, therefore, it would be better to call it a stream of *Sciousness,* a *'Thinker'* thinking objects, some of which it thinks of as 'me', some as 'not-me' but 'not yet including or contemplating its own subjective being' (p 304). The existence of this 'Thinker'

is given as a 'logical postulate' and not a 'direct inner perception of spiritual activity'.

James acknowledges that the hypothetical speculations about *sciousness* 'traverse common sense' and he decides to put them aside. However, the passage, that James modestly dismisses as 'a parenthetical digression', is full of interest. It takes us outside the 'empirical' self. Earlier in this chapter James introduces the 'stream of consciousness' as a 'concrete' existence in time (1890, p 296, cited above). The 'Self of selves' abstracted from it is still concrete enough for James to claim that he can introspect it. Yet, just a little later James is saying that the stream of consciousness is not 'known immediately' (p 304) and suggests that it should be referred to as 'the stream of *Sciousness*'. James acknowledges that this is a 'logical postulate'. Nevertheless the move from discussing empirical data to discussing logical postulates is a swift one. Cooper (2002, p 114) invokes a 'two-levels analysis of the self' to explain apparent inconsistencies between James's position in 'The Principles of Psychology' and in the latter's later writings on the 'non-bodily' self. The two levels are 'scientific' and 'metaphysical' and the distinction already applies to the chapter of 'The Principles of Psychology' that I have been analyzing.

One further remark about James's digression that I will make at this stage is that he claims the speculations at this point 'contradict the fundamental assumption of every philosophic school' (p 304). In fact, they seem to have something in common with theories put forward by the post-Kantian idealists in Germany. Fichte (1797-98/1994) introduces the notion of 'self-positing' to explain without circularity how the 'I' can be conscious of itself. Schelling (1795/1980) postulates as a logical grounding for self-consciousness the 'absolute I', which, like '*Sciousness*', cannot itself be known. Fichte's views will be relevant to my treatment of Nozick in chapter three.

After his 'digression' James puts aside his hypothesis and returns to the 'common-sense' approach, in which 'direct awareness' (1890, p 305) of thinking is assumed. As a result of his own introspection he concludes that 'the part of the innermost Self which is most vividly felt' is a collection of 'cephalic movements'. There is 'an obscurer feeling of something more' but the question of what it is remains unanswered.

Having dealt with the 'constituents' of the self James proceeds to the 'emotions of self' aroused under the three constituent headings. He deals first with 'self-feeling' (called 'self-estimation' in the table). The examples in the table show the sort of emotions he is concerned with here. He identifies 'pride' and 'modesty' as the two 'primary' feelings (under various names). He notes that there is 'a certain average tone of self-feeling' with which the individual identifies, somewhere between the extremes of confidence and its lack, between which we oscillate without rational cause.

James goes on to deal with 'self-seeking', which, he says, refers to the 'fundamental instinctive impulses' (p 307). Among these, the 'material' ones, such as hunger, fear and anger, relate to self-preservation. The 'social' ones are instincts for friendliness, attention-seeking, envy and the desires for power and fame. We select the ones which 'prove serviceable as means to social ends' (p 308). The 'spiritual' impulses are impulses towards intellectual progress, but, James suggests 'much that commonly passes for spiritual self-seeking in this narrow sense is only material and social self-seeking beyond the grave'.

Having reviewed the categories of his table of the empirical self, James considers two issues relating to that self:

1. Choice
James envisages not just a variety of social selves, as indicated above (see the discussion of the 'social' constituent) but a number of material and spiritual

13

selves. We are limited by our physical and intellectual capacities but normally have scope for more than one project or career. To sustain our self-regard we choose goals that are within our reach, if possible, and then are not disheartened by failing to reach the goals for which we do not strive. '*Our thought... here chooses one of many possible selves or characters, and forthwith reckons it no shame to fail in any of those not adopted expressly as its own'* (1890, p 310).

This 'choice' compares with the development of the Nozickian 'metric'. I choose my own personal identity and decide which of my properties a continuer must have in order to count as me. The aim of social recognition is what 'beckons me on' (p 315), at least to begin with. But James sees an enhanced sophistication of choice as the individual gains maturity and loses trust in society's estimation of his achievement. 'For motives of honour and conscience' he may defy social opinion. He then requires an 'ideal judge', God.

As I have indicated, I believe the self develops within a social context, whether in conformity with or reaction against social pressures. When the individual reacts against the society in which he finds himself it may well be because of principles he regards as honourable or godly. But note that, though the tendency to principled behaviour may be in-born, the principles themselves are acquired.

2. Care

James notes that bodily needs are the first object of care, followed by the social need for approval. and finally the spiritual, 'my more phenomenal and perishable powers, my loves and hates, willingnesses and sensibilities, and the like' (p 323). Self-love is not '*love for one's mere principle of conscious identity*'. The mind is dependent on the body 'by an inscrutable necessity'. The warmth of care requires these non-abstract objects.

James's representation of the 'care' or 'personal concerns' (Johnston's phrase, see chapter two) that are a feature of personal identity is compatible with a social interpretation of that concept. The self-love that is expressed in concern for the body is akin to love for and by others and, no doubt, develops (or fails to develop) in concert with the attitudes of the family and friends close to the growing child.

The pure ego

James moves from discussing the empirical self to discussing the 'pure ego', suggesting that the latter is the source of a sense of personal identity and noting that this sense can be treated subjectively or objectively. '...this consciousness of personal sameness may be treated either as a subjective phenomenon or as an objective deliverance, as a feeling, or as a truth' (1890, p 331).   James must mean that it is the sameness, not the consciousness, that can be treated either as a 'feeling' or as a 'truth'. Normally the feeling and the truth converge; exceptionally, my feeling that I am the same person as yesterday might be unjustified. Suppose, for example, I feel today that I am and always have been Napoleon. The feeling would be illusory.

James's concern at this point, however, is to investigate the relationship between the 'subjective phenomenon' and the 'objective deliverance'. He is at pains to point out that the sense of personal identity he is talking about is not the synthesizing subject required for all knowledge, suggested by Kant. It is a sense of identity attributed in thinking to the objects of thought, 'a present self and a self of yesterday' (p 332), both felt with a 'warmth and intimacy' which marks them as ours. We have come back to the idea of a stream of consciousness, with thought itself as the subject 'Thinker' relating at each moment (segment of the stream) to the past and present objects of consciousness.. To clarify his conception of this relationship James uses the metaphor of a herd of cattle and their owner. '...the owner picks out and sorts together when the time for the round-up comes in the spring, all the beasts on which he finds his own particular brand' (pp 333-334). The process of recognition and, to use James's word, of

'appropriation', for which this is a robust metaphor, is a delicate mental one. Yet, the metaphor makes sense only in a social context. The cattle only need to be branded because there are other cattle around, and the branding system has to be agreed with the other owners. This is not to say that there is a danger of confusing the objects of one individual`s thought with those of other people`s thought. The metaphor cannot be taken too far. But James is right that the individual has a special feeling, reasonably designated `warmth`, towards his own objects of thought, as to his own body. If the warmth is lacking, the sense of identity is lacking. This sense of identity is achieved through social interaction (cf. Mead, 1934, discussed in the next chapter). It is in that way that the branding is a social phenomenon

While not emphasizing the social nature of this ownership, James does emphasize its 'reality'. The cattle are branded, he maintains, because they *really* belong to the herd. The selection is not arbitrary. It is common sense that 'the unity of all the selves' represents a 'real belonging' to a what James calls a 'pure spiritual entity of some kind' (1890, p 337). The owner of the cattle is 'the real, present onlooking, remembering, "judging thought" or identifying "section" of the stream' (p 338). This 'Thought', which James now spells with a capital 'T', is a 'pulse of cognitive consciousness' which inherits the title of ownership from the immediately preceding Thought, together with that Thought and all its contents

The 'Thought', then, is what in modern parlance we might call the present 'time-slice' of the stream of consciousness, the 'present self', immediately overtaken by another. But for James the present Thought is only a subject of consciousness and does not become an object of consciousness until it is past and the next Thought has taken over as the 'present self'. He introduces another metaphor, the 'chain', to elucidate his conception of the relationship between the subject 'Thought' and the objects of consciousness. '... the Thought' he says 'never is an object in its own hands, it never appropriates or disowns itself. It appropriates *to* itself, it is ... the hook from which the chain of past selves dangles' (p 340, James's emphasis).

It is in this way that James avoids the circularity that worried Fichte and other post-Kantians. (See the paragraph on Fichte in chapter three.) James's 'judging Thought' is exclusively a subject 'hook' and is not itself observed until it becomes a link in the chain suspended from another hook.

Though James does not highlight them, there are similarities between the 'Thought' described here and the 'Thinker' of the 'Stream of *Sciousness*' hypothesized earlier in the chapter, not 'thinking its own existence'. James emphasizes the 'reality' of the Thought and its appeal to the common-sense view of the unity of the self over time (see p 338). He considers that the notion of appropriating the previous Thought and all its contents deals with this 'urgent demand of common-sense'. Yet, as I have said, he calls it a 'spiritual entity' and 'the Thought never is an object in its own hands' (p 340). It seems that it cannot be an object of consciousness until it is past and taken over by the next Thought. It seems to hover on the edge of the 'metaphysical level' inhabited by the 'Thinker'.

James compares his theory of the self with the 'Spiritualist' theory (of the soul), the 'Associationist' theory (of distinct ideas relating to each other) and the 'Transcendentalist' theory (p 342 f). Of the concept of the 'soul' James says that it is 'needless' for explaining the 'subjective phenomena of consciousness', which can be satisfactorily explained by his own theory of the 'stream' (p 344). But James's rejection of this idea of an *immaterial* agent, at least at this stage in his writing, goes beyond regarding it as unnecessary. Though he acknowledges that there is 'something more' which connects thoughts with brain states, the 'soul' concept does not provide the connection. In this context 'soul' is 'an illusory term' (p 346).

The 'substantialist' view of the soul is associated by James with Descartes, Locke, Hobbes and Leibniz (1890, p 344). The 'Associationist' theory he ascribes to Hume, whose *Treatise on human nature* he quotes at length (p 351). Hume

writes: 'I never can catch *myself* at any time without a perception' and concluded that the self is 'a bundle or collection of different perceptions' (1738/1911, p 239). We attribute identity to objects that are merely related by 'resemblance' (p 240). It is not so much that Hume denies the 'thread of resemblance' connecting the perceptions, although James says he does (1890, p 352). What Hume denies is our ability to perceive the reality of this connection. (See Appendix to 1738/1911 book I, quoted in James 1890, p352.) For James, the 'stream of consciousness' provides a 'real' connection through the unity of its parts (p 353).

James says that according to the transcendentalist Kant the self (the *I think*) is the necessary precondition for all consciousness, but we have no conception of what the self is. (See James 1890, p 362.) Kant and his successors, James claims, do not distinguish the Ego's relation to the objects of which it is aware from its awareness of itself and of it separation from the objects. 'That the Object must be known to something which *thinks*, and that it must be known to something which *thinks* that *it thinks,* are treated by them as identical necessities' (1890, p 361 note). James says Kant only identifies the Ego as the subject of experience and denies knowledge of the 'soul'. James is right that Kant does not attempt to analyse the self independently of experience[3], but at least some of Kant's successors in Germany were aware of the problems of self-consciousness not solved by Kant (see Fichte, 1797-98/1994, Schelling, 1795/1980). To clarify his own view James distinguishes between 'the objective person', which is known, and the 'passing subjective Thought' which knows it. '*Hereafter let* us *use the words* ME *and* I *for the empirical person and the judging Thought.*' (p 371) The 'me' changes over time and the 'I' finds only a 'relative' identity in it (p 372). 'Thus the identity found by the *I* in its *me* is only a loosely construed thing, an identity "on the whole", just like that which any outside observer might find in the same assemblage of facts'.

---

[3] Eg *Prolegomena* Footnote 3 to §46: 'Were the representation of the apperception (the Ego) a concept, by which anything could be thought, it could be used as a predicate of other things or contain predicates in itself. But it is nothing more than the feeling of an existence without the least definite conception and is only the representation to which all thinking stands in relation' (Kant 1783/1977, p 75).

James's treatment of personal identity here contrasts a little bit with his treatment earlier on, when he is emphasizing its factuality. 'The intellectual operations seem essentially alike,' he suggests (1890, p 331), 'whether I say "I am the same", or whether I say "the pen is the same, as yesterday"'. Yet, there is a difficulty in explaining what counts as 'identity' for organic things, that does not apply to inanimate objects. Locke (1689/1975, pp 330-331) discusses the identity of the oak which develops from an acorn. James is certainly aware of the changes that the empirical 'me' undergoes. He is also aware, as the preceding paragraph shows, that the 'outside observer' takes account of such changes when attributing personal identity.

James selects two types of change in the 'me' for discussion at the end of the chapter:

1. **'Alterations of memory'**

James includes in this category 'false' memories, which, as he points out, are common. We embroider experiences in the telling and then remember what we narrated rather than the original experience. I would argue that something similar happens to stories about ourselves we learn from others. These may be true or untrue but we often accept them and they can come to seem like our own memories. James denies this. He writes (1890, p 335): 'We hear from our parents various anecdotes about our infant years, but we do not appropriate them as we do our own memories.' Whatever the truth of this particular point, it is clear that 'false' memories generated by social circumstances impact on what we claim as our personal identity.

James also includes 'loss of memory' in the category of 'alterations of memory'. James notes that in dreams we forget waking life; on waking we forget the dreams. 'Should a man be punished for what he did in his childhood and no longer remembers?' (p 373). Locke also writes about this problem: 'Socrates waking and sleeping is not the same person' (1689/1975, p 342). Locke would

answer 'no' to James's question. Yet what we can remember is not necessarily the same as that for which we can be held responsible (can be proved to have done).

2. 'Alterations in the present bodily and spiritual selves'

James characterizes these as 'abnormal' and of three types: insane delusions; alternating selves; mediumships or possessions. He narrates a number of cases to illustrate each class. He does not attach importance to the classification, because we have little knowledge of the 'elements and causes' of these occurrences (1890, p 375). But he points out that individual may or may not be aware of the changes. In one case, a patient of Dr Krishaber (p 377 f), the man suffered extensive delusions, but was aware of them as such and could write of them afterwards. James claims: 'In cases similar to this, it is as certain that the I is unaltered as that the me is changed' (p 378). In another case, that of Mary Reynold cited by Dr Weir Mitchell (p 381 f), the individual woke up after a long sleep with totally different personality traits and with no memory of her previous state. Some weeks later she reverted to her previous self and then continued alternating states for about 15 years. James does not say as much, but, presumably, he thought that in such cases, the 'I' as well as the 'me' is altered.

Conclusion and ways forward

James concludes that 'the same brain may subserve many conscious selves, either alternate or coexisting' (1890, p 401). At present society does not come to the same conclusion. At the level of the 'empirical life of the Self' James's account of the 'material' and 'social' constituents conform to the common-sense view of the self, identifying one personal identity with one person's life. Different 'social selves' are regarded as different roles of the same person. The person playing the roles is recognized as one person and is the object of care, blame etc.   There is a distinction to be made between different kinds of 'social self'. The empirical social roles, such as those adopted at work in contrast with those adopted at play, are taken up and put down at will. Between divided selves, such as those of Mary

Reynold and those of Dr Jekyll and Mr Hyde (considered in the next chapter), the switch is not so easily achieved. Common-sense concerns, such as the requirements of etiquette and the imposition of laws, take account of these differences. However, thought experiments involving the transplant of brains or brain-contents to different bodies, to be discussed in later chapters, suggest the common-sense view may be inadequate.

A person undergoing the sort of changes of 'me' that James is discussing in the last part of his chapter is also regarded as one person, even if he is regarded as an abnormal or sick person. Such abnormalities cause problems for the individual who does not remember in one state what he did in another state. They also cause problems for society. Can you blame Mary Reynold in either of her two 'states' for what she did in the other? If such cases were more common society might have to find a different way of dealing with them.

The 'stream of consciousness' is James's great contribution to the discussion of the 'self'. With it he advances beyond a tabulation of personal characteristics towards a psychological and metaphysical account of what the 'self' is. As mentioned above, James uses the 'stream of consciousness' concept to justify his rejection of the 'substantialist' view of the soul. He also sees it as overcoming the 'transcendalist' failure to distinguish between the self as the subject of consciousness (with regard to objects external to the mind) and the self as the object of self-consciousness. Furthermore, the notion of the 'Thought' appropriating all previous thoughts provides a view of continuity as something more than just 'memory'. Strawson reads James as claiming that the 'Thoughts' are 'numerically distinct pulses of thought, riding on continuous brain-changes', 'actually existing physical entities' (2009, p 558). Strawson develops some of the metaphysical implications of this approach. He proposes that these 'subjects of experience' are the best candidates 'for being physical objects or substances'.

Cooper (2008b, p104) compares James's view of the self with Nozick's. 'Nozick's account of entification, including the decision value refinement of it, is doing the same work as James's "legal title" handed from one herdsman to the next'. Certainly Nozick's 'reflexive self-referring act' synthesizing the self in a moment has something in common with James's Thought appropriating all previous thoughts. But the Thought is not self-conscious; the *Thinker* of the stream of '*sciousness*' is not 'including or contemplating its own subjective being' (1890, p 304). while Nozick's reflexive self-referring act is paradigmatically self-conscious.

Nozick gives the impression (to be elaborated in chapter three) that the individual actively chooses who he will be. James's picture is more of an individual finding himself and, I would say, finding himself in a social context. James, at least attends more to the social elements of personal identity, such as the social drives which are not just learnt responses but 'fundamental instinctive impulses' (1890, p 307). It is true that from the social perspective James's focus is on empirical factors such as the varying roles an individual plays and the urge for recognition by others. He does not highlight the social ingredients of the 'Pure Ego', the sense of self which is both subjective and objective (p 331, quoted above). Mead does do that, as we shall see in the next chapter. It is these social dimensions of the personal identity picture, both 'empirical' and 'pure', that I seek to emphasize.

CHAPTER 2: OPEN TO THE PUBLIC

Introduction

In this chapter I will consider the views of philosophers who have paid attention to the 'social elements' of personal identity. 'Paying attention' to these elements does not mean ascribing to them total power or even dominance. But it does mean not neglecting them. I count William James as one such philosopher, although, after my extended treatment of him in chapter one, my treatment of him here is brief. I also include George Mead, Mark Johnston and Eric Olson. My treatment of Olson is also brief, intended to do little more than introduce the idea of 'biological continuity'.

It will be helpful to bear in mind one general factor: the 'levels' of interaction between society and the individual. There are levels at which the interaction is obvious, as with James's 'social self', which he describes as the 'recognition' (1890, p 293) bestowed by society. Conscious requirements come from different social groups which the individual consciously meets or rejects. There are other levels at which social influence is less obvious, as with beliefs and emotions. They are the individual's property; their social sources may be rarely thought of or altogether forgotten. Yet personal identity is a human phenomenon and 'when we refer to human nature, we are referring to something which is essentially social' (Mead, 1934, p 139, note 2). At the end of the last chapter I differentiated between the social ingredients of the 'Pure Ego' and those of the 'empirical self'. There may well be correspondence between 'levels' of social interaction and 'levels' of the self. All I am claiming here, however, is that there is social influence at one level or another present at each stage of the development of personal identity; where it is not immediately perceptible, it still lurks below the surface.

James

Among the philosophers to be considered in this chapter, I will deal first with
James and elaborate on issues from his analysis relevant to my thesis. I have
already referred to the empirical 'social self'. Potentially, according to James, a
subject has a different 'social self' for each other individual or group of
individuals in his life. Multiplicity of social roles is a familiar phenomenon.
Most people are aware in themselves and in others of the tendency to behave
differently in different environments. Normally this portends nothing more than a
healthy adaptability. When Shakespeare's Henry V called upon his soldiers to
'disguise fair nature with hard-favour'd rage' ('Henry the Fifth', Act III, Scene I),
he was appropriately asking them to apply to their role as soldiers qualities that
they would not associate with their roles as husbands or fathers. Such a change of
role does not imply a change of personal identity. The individual would not fail
to recognize himself or fail to be recognized by others when he had moved from
one role to another. The two roles would certainly feel different; there would be
psychological differences. There would also, perhaps, be physical differences –
differences of clothes, differences even of physical bearing. But all these would
not amount to discontinuities.

A point to note about James's social roles (and Henry V's soldiers) is that there is
what might be termed a 'social-contract' element involved. The individual
adjusts to what he perceives as society's requirements and the people around him
make these requirements clear. This happens in a range of ways. A soldier or
doctor receives extensive training; the roles make big demands on those who take
them on and society's 'recognition' is commensurate. Yet, the supposedly less
demanding roles of father or husband also have social aspects, and not just
because to be either you need someone else's co-operation. The child absorbs
ideas of what it is to be a parent and a partner, if not with its mother's milk, at

least with its mother's, and father's, care, and, later, with awareness of the expectations of people outside the family group[4].

In cases like that of Mary Reynolds (James, 1890, p381 f) the 'contract' is not consistently effective. For many years her life alternated between two totally different personae, without continuous memory. That is to say there was no psychological continuity, though there was physical continuity. Since there was physical continuity society treated her as one person. Unlike the role changes James considers as an effect of the 'social self', these changes of person were uncontrolled and not required by the society around the woman, which must have found them a source of great embarrassment. However, she does not appear to have acted criminally even in her more aggressive role, so society was able to cope with her. People who are nowadays labelled as 'schizophrenic', and at certain times act criminally, are controlled by society with the use of drugs or some form of detention. A personal identity in breach of the social contract must be suppressed in some way. This type of event is fictionalized in R.L.Stevenson's 'Dr Jekyll and Mr Hyde', where the two personae are presented as two sides of the same person and the metamorphoses are to some extent controlled by the subject, using a potion. Mead, whose views I am about to discuss, describes the phenomenon this way (1934, p144):

> The phenomenon of dissociation of personality is caused by a breaking up of the complete, unitary self into the component selves of which it is composed, and which respectively correspond to different aspects of the social process in which the person is involved, and within which his complete or unitary self has arisen; these aspects being the different social groups to which he belongs within that process.

The anti-social individual is no less a product of society than the well-integrated. Edward Hyde, rebelling against the restrictive Victorian society in which he lives,

---

[4] See references to Mead 1934 below.

is comparable to the drug addicts of today, who form their own alienated, sometimes criminal, groups in which they can retain personal identities unacceptable to the established community. Abnormally the different (socially evolved) personae of one individual harden into what might be thought of as different personal identities. In this abnormal sort of case the individual in one persona is not able to control what he does in the others, is sometimes unconscious of the others. If such cases were more frequent, society might find it appropriate to treat them as different persons within the same body.

I showed in chapter one how James identifies the 'spiritual self' with the 'stream of consciousness' and, in particular, with a segment of that stream, the present 'Thought', which appropriates all (and only) the previous Thoughts in that stream. This he says is the 'concrete' view of the spiritual self, in contrast with the 'psychic faculties or dispositions' which constitute the 'abstract' view (1890, p 296). A little later (p 304) he talks of the 'two parts' of objective experience, 'Self' and 'not-Self', beyond which 'there *is* nothing save the fact that they are known, the fact of the stream of thought being there as the indispensible subjective condition of their being experienced at all' (James's emphasis). This is the moment when James considers calling the stream of consciousness a stream of '*sciousness*', each bit of which has its 'me' and its 'not-me' but is not yet 'including or contemplating its own subjective being'. James does not develop this metaphysical 'logical postulate' very far, but it is fair to present it as a logical postulate for empirical experience of the environment, in contrast with the logical postulates of the idealists Fichte and Schelling[5]. The 'not-me' for James is a reality outside the self of which the self can be conscious and which forms part of the environment in which personal identity develops.

---

[5] James belongs to the American 'pragmatism' movement. I suspect pragmatism implies empiricism. At any rate, it is clear that James subscribes to both bodies of belief. He defines 'pragmatism' as the idea that the truth of propositions or beliefs 'consists in certain definable relations between them and the reality of which they make report' (1907/1982, p 134). Idealists believe that truth is limited to what is in the mind.

Mead, 1934

For James's fellow-pragmatist Mead, society is intrinsic to the very constitution of the self, to what James describes as 'the Pure Ego'. 'The self, as that which can be an object to itself, is essentially a social structure, and it arises in social experience' (Mead, 1934, p 140). On Mead's account, someone does not become a person until he can interact with the society around him and that interaction is essential to being a person. 'The self is something which has a development; it is not initially there, at birth, but arises in the process of social experience and activity, that is, develops in the given individual as a result of his relations to that process as a whole and to other individuals within that process' (p 135). Mead differentiates two stages in this development, the first in which 'the individual's self is constituted simply by an organization of the particular attitudes of other individuals toward himself and toward one another in the specific social acts in which he participates with them'; the second in which 'that self is constituted not only by an organization of these particular individual attitudes, but also by an organization of the social attitudes of the generalized other or the social group as a whole to which he belongs' (p 158). The 'generalized other' is 'the organized community or social group which gives to the individual his unity of self' (p 154). Mead is differentiating two processes: in the first, the individual, presumably as a young child and therefore largely unconsciously, develops his own personality by absorbing the traits of those around him. In the second process, society actively, though, again, not necessarily consciously, imposes characteristics on the individual. The kindergarten, for example, 'takes the characters of these various vague beings and gets them into such an organized social relationship to each other that they build up the character of the little child' (p 153). Probably the two processes overlap. Certainly, in the second process, some adults consciously instil what they regard as appropriate traits. Parents and teachers are obvious examples of intentional educators. Other members of the organized community achieve their results less deliberately by means of the immature individual's tendency to imitate. In any case, the developing personal identity is flexible or, as

27

Mead has it, 'vague'. Evolution into a mature individual is a social process. The previous paragraph considered some of the results when the social pressures are too great or have a disrupting effect.

In the context of my general thesis, it is interesting to compare Mead's account of how the self arises with Nozick's suggestion that the self is 'synthesized around the act of reflexive self-referring' (Nozick, 1981, p 87). For Mead, too, 'the language process is essential for the development of the self' (Mead, 1934, p 135). For Mead, as for Nozick, there is significance in how the word 'self' is used. Mead says it is 'a reflexive, and indicates that which can be both subject and object' (p 137). Certainly, in English, the word as a suffix can be used as both subject and object, as in: 'I myself see myself'. But this pattern is not the same for all languages. Compare, for example, the French: 'Je me vois moi-même'. The philosophical significance of the linguistic practice is limited. Probably what matters, both for Nozick and for Mead, is the intellectual development accompanying the linguistic skills. In so far as the individual is fully aware of himself as a person when he refers to himself he has achieved personal identity.

According to Mead, the individual becomes a 'self' 'not directly or immediately, not by becoming a subject to himself, but only in so far as he first becomes an object to himself ... by taking the attitudes of other individuals toward himself within a social environment or context of experience and behaviour in which both he and they are involved' (p 138). In other words, the child has to mature sufficiently to see himself at least to some extent objectively, which he learns to do from noticing how others react to him on social occasions in which both he and they are involved. One might say that Nozick makes the opposite claim: the individual achieves personhood by becoming a 'subject to himself', when he has the capacity for *reflexive self-reference*. (This is explained in chapter three.) I would like to offer a resolution to this hypothetical argument between Nozick and Mead by suggesting the individual becomes *both* an object and a subject to

himself. Consciousness (as a subject) involves self-consciousness (as an object). Both elements develop in sophistication up to and beyond the point of personal identity. This is a suggestion developed in chapter 4.

 Mead distinguishes this 'objective' self, which he refers to as the 'me', from the 'subjective' self, the 'I'. 'The "I" reacts to the self which arises through the taking of the attitudes of others', that is, it reacts to the "me" (p 174). The distinction, he claims, is not a 'metaphysical one'. 'I do not mean to raise the metaphysical question of how a person can be both "I" and "me"'(p 173). But the distinction is a reflection of human 'conduct'. A comparison with James is illuminating. According to Mead, James looks for 'the basis of the self in reflexive affective experiences, i.e., experiences involving "self-feeling"'. But this 'does not account for the origin of the self, or of the self-feeling which is supposed to characterize such experiences' (p173). Mead, as we have seen, is convinced that there is a social basis for both the self and self-feeling. Nevertheless, there is a parallel between Mead's account of the "I" and the "me" and James's model of the 'chain' linking the current self, the 'Thought', to past selves (see chapter one). Mead writes (p 174): 'The "I" of this moment is present in the "me" of the next moment...the "I" in memory is there as the spokesman of the self of the second, or minute, or day ago. As given, it is a "me," but it is a "me" which was the "I" at the earlier time.' Similarly, James's current Thought 'never is an object in its own hands' but is the hook holding the chain of past selves (Thoughts become objective), (James 1890, p 340).   An advantage of Mead's 'chain' is that it links the present moment to the future as well as the past. The psychological continuity which underpins personal identity involves anticipation as well as memory.

One more point about Mead (1934), before this chapter moves on from him. From the arguments I have already set out, he concludes: 'one has to be a member of a community to be a self' (p 162). Yet this may seem a surprising notion against the background of normal use of the word 'self'. Mead himself acknowledges this. 'It has been the tendency of psychology to deal with the self

as a more or less isolated and independent element, a sort of entity that could conceivably exist by itself. It is possible that there might be a single self in the universe if we start off by identifying the self with a certain feeling-consciousness' (p 165). A living body with conscious states could exist in isolation and be called a self. But Mead's point is that this would be a different sort of self from one developed in the social context he sees as crucial to self development as we know it. A comparison might be made with moral agency. James (1891) envisages '*moral solitude',* a universe with one 'sentient' being, for whom there would be 'good' in so far as he made things good for himself but 'no outward obligation'. Richard Taylor (1984), conducting a similar thought experiment, points out that, though 'good' and 'evil' could be said to exist in a world with just one person in it, 'right' and 'wrong' do not exist until there is more than one moral agent in the world. A self could exist in isolation. Yet such a self would neither need nor be able to develop the qualities associated with personal identity as the term is used in this paper. What Johnston (1989) calls 'person-directed concerns' – these are explained below   - would apply either not at all or in such an attenuated form as to be unrecognizable. The lone individual would have concerns and expectations relating only to himself. In so far as he had a concept of 'his life' it would be different from that engendered by the perception of other lives or consideration of the effects on others of his own life.

Olson, 1997
Mead, while emphasizing the social input into the beginning and maintaining continuity of personal identity, treats it as a psycho-physical phenomenon. Mental events, such as thoughts, feelings, attitudes and emotions feature strongly in his account of the self. Yet, so, too, do physical events. He sees play as important to the development of the self. (See, e.g., Mead, 1934, Section 20 *Play, the game, and the generalized other*.) 'The relationship of the "me" to the "I" is the relationship of a situation to the organism' (p 280). It is the central nervous system and the physiological capacities of this 'organism', 'the human animal', that enable it to develop a self (see p 237 and note, p 236).

However, there is a potential dichotomy between physical and psychological continuity, which is a feature of Olson's position. It is Olson I am going to discuss in this section, but, before doing so, I shall fill in a little of the historical background. Descartes is notorious for his theory of 'dualism': the notion that mental substance is separate from and independent of material substance. 'It is certain', he says (1641/1993, p 51), 'that I am really distinct from my body and can exist without it.' Using modern phraseology, he would say psychological continuity is what matters for the self and that is possible without physical continuity. Locke thinks being the same 'substance' is different from being the same 'person' and that, again, is different from being the same 'man' (1689/1975, p 332). Being the same man consists in 'participation of the same continued life' (p 331). Someone will be 'the same *self* as far as the same consciousness can extend to actions past or to come' (p 336). So Locke, like Descartes, associates continuity of personal identity with psychological continuity but, unlike Descartes, dissociates it from continuity of substance. This is illustrated in Locke's story (p 340) of the 'soul' of a prince entering the body of a cobbler when the latter is 'deserted by his own soul'. The resultant being is, according to Locke, the same 'person' as the prince, having the consciousness of the prince's past life and being 'accountable only for the prince's actions'. Yet the 'man' is determined by the body and the being would be 'the same cobbler to everyone beside himself'.

Olson (1997) claims that there are two approaches to personal identity, that of 'psychological continuity' and that of 'biological continuity'. The former approach asserts that continuity is determined by facts about memory, character and other mental features. The latter approach suggests it is determined by biological functions such as metabolism, respiratory mechanisms and blood circulation. What counts is that these functions persist, and biological continuity can be maintained without psychological continuity. A human whose cerebrum has been removed can survive, with sufficient brain-stem to regulate the

biological functions (and with a great deal of intervention from the medical society around him). Olson's argument is that we are essentially human animals. 'Animal' is a 'substance concept' (p 30), whereas 'person' is a 'phase sortal'. That is to say that a human being exists if and only if he is a human animal but can exist without being a person, just as he can exist without being a child, 'child' being another 'phase sortal'. ('Substance' and 'phase' sortals are discussed below in the section on Johnston.) The human animal can persist in a vegetative state without a cerebrum, but such an animal can no longer be described as a 'person'.

In Olson's (1997) version of Locke's story of the Prince and the Cobbler, Prince's cerebrum is transplanted to Cobbler's body. Prince's body survives as 'Brainless'. Cobbler's body with the Prince's cerebrum becomes 'Brainy'. Brainy, with the memories and characteristics of Prince, might think of himself as Prince but would undoubtedly feel strange in his new body and, perhaps, even stranger when he becomes aware of what exactly has happened. According to Olson (p 65), Prince's family and friends would accept Brainy as Prince. Cobbler's wife would not. I think we cannot be sure, but we can easily envisage disagreement in the society affected about which person has survived. Brainless does not have a vote. Is it up to Brainy to make the decision about whether he is Prince? Can he disregard the views of Prince's and Cobbler's family? The Nozickean approach that I shall be discussing in the next chapter might indicate a positive answer to these questions. For the moment, it is enough to say there could be different views.

Johnston, 1987
Johnston (1987, p 61), defines '*wide* psychological continuity' as the view that 'mental continuity and connectedness can constitute personal identity even if the holding of these relations is not secured by its normal cause, the persistence of a particular human body or brain'. This accords with Olson's understanding of 'psychological continuity'. Johnston (p 75f) discusses Sydney Shoemaker's imaginary case in which Brown's brain is transplanted to Robinson's head, a case

similar to that of Olson's Brainless and Brainy. Johnston notes that the 'predominant reaction' is to suppose that 'Brownson' (Brown's brain in Robinson's body) is Brown, but argues this 'predominant reaction' is generated by the 'distorting influences' of 'psychological- and social-continuer effects' (p 77). The article explains these 'effects' a little later (p 82) as the interest in retaining 'some occupant of the particular complex of detailed and manifold social roles (father, lover, friend, leader, supporter, colleague, nemesis, regular customer, etc.) which made up that person's social life'.

Johnston (1987), then, conceives a multiplicity of social roles as empirical elements of the self. They are not logically necessary (for Johnston) in the sense that to have the same personal identity means to have the same social roles. In this paper he regards them as having a 'distorting effect'.   'The wide psychological view is thus parasitic upon the psychological- and social-continuer effects...persons antedate, outlive, and may sometimes be outlived by their personas' (p 82). However, a 'distorting effect' is still an effect. There are external pressures to sustain social roles as part of the personal identity, as on Henry V's soldiers. There are internal pressures such as *habit* (James's term discussed in chapter 3). Johnston considers that thought experiments like Shoemaker's and Bernard Williams's[6] (and, presumably, Olson's) provide unreliable evidence of how the concept of personal identity is interpreted. For us, who are 'essentially human beings', he says, there is 'nothing problematic in our everyday practice of re-identifying ourselves and others on the strength of the continuous mental and physical functioning of the human organisms we encounter' (p 83). Problems do arise in rare cases, like that of Mary Reynolds. In such cases it seems that society's view, that there has to be one continuous personal identity because there is one continuous body, is given priority over the

---

[6] Williams (1976/1970) in an influential paper envisages an 'exchange of bodies' between two persons A and B, in the sense that A's mental contents are transferred to B's body and vice versa. Before the exchange each is offered a reward or punishment to the other's body. Telling the story one way, Williams makes it seem as though A before the operation would be more concerned about what happens to the 'B-body' after the operation. Telling it another way, he makes it seem as though A would be more concerned about what happens to the 'A-body'.

individual's view. Though each of Mary's personae would have thought of herself as one person, neither would have felt continuous with the other. If problems arose routinely, as a result, say, of frequent transplants of successful brains into successful bodies, would the one-body/one-personal identity view still prevail?

Johnston, 1989

The answer to that last question might be teased out of Johnston's later, longer and more elaborate paper on personal identity, 'Relativism and the Self' (1989). Here he develops a relativist approach to personal identity, maintaining the thesis that the concept of personal identity is a 'dependent justifier'. He defends what he calls the 'surprising view' that 'what makes some process count as my surviving is...a matter of what we are inclined to believe' (p 441).   What makes this view surprising is that we are used to thinking of survival as a matter of fact. The belief that a subject has survived or not is, we assume, justified by evidence of a familiar kind, very much the same sort of evidence we would accept for a judgement that he was alive or dead. However, Johnston points out that the concept of personal identity is indeterminate in a way that the concept of being alive is not. He uses Parfit's thought experiment envisaging a machine which transforms one person into a completely different person in very small stages (Parfit, 1984, p 231 f). It is not possible to determine a particular point at which the person who enters the machine has changed personal identity. The 'sorites' problem is well known and it does not particularly worry us that we cannot say exactly how many hairs you have to lose before you become bald. But the application of the 'sorites' model to personal identity might stimulate reflection. Might Dr Jekyll have modified his dose of potion so that he could have satisfied his wilder urges without so obviously becoming Mr Hyde?

Johnston (1989), at any rate, finds the idea helpful in its application to psychological (as well as physical) continuities. The indeterminacy involved, he suggests, can be treated as 'semantic' (p 448). For situations in which semantic

indeterminacy matters we are used to making arbitrary distinctions. We could, for example, arbitrarily stipulate what counts as bald for the purposes of choosing a model for a hair-restorer advertisement. But in respect of personal identity, as Johnston points out, there are what he calls 'person-directed concerns'. He explains these (p 448) as:

> '(i) one's future oriented and retrospective concerns for oneself and others;
> (ii) one's expectations about experiences and memories of those experiences;
> (iii) one's expectations about the relations between action and desert.'

This listing is an elaboration of what Gendler (2009), quoted in chapter one, describes as the 'practical and emotional significance' of personal identity. I emphasize again that these concerns apply to the society interacting with the individual as well as to the individual. Johnston (1989) stresses that the indeterminacy of personal identity causes problems in relation to these concerns. Precisification might run in the face of them, since it could be different for different languages. A speaker of one language might regard himself as surviving an event that the other would take to be death, simply because they 'internalized different stipulations about what previously counted as indeterminate cases' (p 451). Someone with a bald patch might be reckoned 'bald' as a member of a hairy hippy tribe and 'not bald' in a group of shaven punks.

To clarify his conception of the problem Johnston (1989) imagines a situation in which 'teletransportation' is available, a means by which a subject enters a machine which destroys him and replaces him with an exact psychological and physical replica in another place. The replica differs only in having a completely healthy body. Johnston also imagines two communities: 'Human Beings' for whom teletransportation is suicide and replacement by a duplicate; 'Teletransporters' for whom it is a fast means of travel, with the added benefit of curing all diseases and disabilities. In Johnston's story some of the Human

Beings come to see the advantages of the Teletransporters' view of personal identity. These 'relativist' Human Beings would believe that their only difference from the Teletransporters was that they had internalized different stipulations about personal identity. They would attribute beliefs about personal identity to the 'accidents of conceptual history' (p 456). There is no doubt that conceptual history plays a big role in personal concerns and ideas about continuity. The effects of religious beliefs are an example and the discussion of personal identity by Descartes and Locke shows marked differences from the discussion by philosophers of the last and present centuries. In chapter four I speculate about ways in which the conceptual future might affect the idea of personal identity.

Johnston's (1989) answer to my question about what would happen if problem cases such as Mary Reynolds' were more frequent would, I suggest, be that the concept of personal identity would change. Does this mean that conceptual history is purely accidental and that a totally relativist approach to personal identity is justified? Johnston is at pains to try to identify what lies beneath the concept of 'person', which itself lies beneath the concept of personal identity. Referring to work by Hirsch, 1982, and Wiggins, 1967, Johnston suggests (p 462) that being an individual person is to be classified as belonging to a particular 'substance sortal'. This is to say 'person' is a 'term such that necessarily there is no time at which the individual exists without satisfying it at that time' (p462). 'Substance sortals' are contrasted with 'phase sortals', such as 'child' or 'philosopher', through which a 'person' may pass while remaining a person. (Olson, as mentioned above, came to the conclusion that 'animal' is a 'substance concept' and 'person' is a 'phase sortal'.) Johnston adapts the idea to accommodate the Teletransporters' view that a 'person' can continue through different stages by being replicated at the end of each stage. He uses the term *series human-being* (p 456) to describe what someone is on the Teletransporters' account. He suggests that persons are *cryptophase* kinds. 'The relativist should say that kinds like Human Beings and Series Human Beings are, in fact, *cryptophase kinds*, associated with phases which persons can live through given,

but only given, the special circumstances that is (*sic*) refiguration' (p 464). 'Refiguration' is defined (p 463) as a change in the personal identity relation which control's someone's person-directed concerns. Which 'cryptophase kind we fall under depends on our 'person-directed concerns' at the time (p 465).

Before leaving Johnston (1989), there are two issues that I wish to raise (without trying to settle):

1. Johnston can be interpreted as meaning that what counts as personal identity depends entirely on what the individual or the community believes, regardless of the evidence. 'P's being conclusive evidence for q merely is its being the case that we are standardly disposed to believe q in consequence of believing p' (p 441). Yet, his story is convincing only if the teletransportation machine works, at least most of the time. If the Teletransporters find it does not work, the boot will be on the other foot. At least some of the Teletransporters will be attracted to the view maintained by the 'absolutist' Human Beings who consider that personal identity is restricted to one lifetime. Johnston's 'Human Beings' community is postulated as contemporaneous with teletransportation. In the current situation persons only achieve the kinds of continuity which justify person-directed concerns within the confines of one lifetime. Disregarding the qualifications required for childhood and senility and for the abnormal 'multiple-personality' cases, most people can remember and plan for a lifetime and society accepts that span as the basis of its expectations, rewards and punishments. That is to say there is a practical, empirical basis for the concept of personal identity which is agreed between the individual and his society. To say the concept of personal identity is relative to the society under consideration is not to say that it is totally arbitrary.

2. Johnston distinguishes:
'dependent' justifiers, dependent on our practices and 'dispositions to take certain things as justifications' (1989, p 442)

'derivative' justifiers which 'justify preferences only because of the typical concomitants of those facts' (p 443).

Personal identity for him is a 'dependent' justifier of responses such as person-directed concerns, dependent on, for example, the disposition to take the organic and psychological continuity achieved by teletransportation as justification for these concerns. But he also claims that personal identity is a 'non-derivative' justifier. Members of a community accepting a particular version of personal identity would appeal to it in its own right to justify a set of responses, not just because it was 'the manifestation, cause or necessary condition' of other justifiers of those responses (p 458). Thus 'Human Beings' who did not believe that teletransportation continued personal identity would resist it, even if it involved perfect psychological and physical replication.

Yet (p 459) he expresses doubt about having tackled the 'real worry' that 'dependent' justifiers do not provide 'firm ground' for our responses. In so far as this is a 'general' worry he suggests that the solution might be that there could be no independent justifier for these responses, though he does not develop that argument in this paper. Elsewhere Johnston uses the phrase 'response dependent' to describe concepts manifesting 'conceptual dependence on or interdependence with our responses in certain specified conditions' (1989a, p 145). Our concepts of secondary qualities such as colours are examples. In *Relativism and the self* Johnston points out that our concept of redness depends on how certain objects appear to normally-sighted human beings under normal conditions (1989b, p 458). Red objects do have physical properties which cause us to see them as red but the concept is dependent on how we respond; other beings could respond differently. Personal identity is also a 'response-dependent' concept. A replica might be a perfect psychological and physical duplicate, but whether or not it would be seen by a particular group as continuity of personal identity would depend on the group's standards. If a concept is 'response-dependent', there

could be no *independent* justification for its application. This appears to be Johnston's solution to the 'general' worry.

Johnston (1989b) also identifies a 'particular' worry. Using the term 'reculturation' to signify the change in attitude required to accept teletransportation as a means of continuing with the same personal identity, he envisages a situation in which the Teletransporters only provide reculturation at the same time as teletransportation. In that situation a Human Being who had not already been recultured would not be justified in accepting teletransportation. Johnston thinks that, if the justification or lack of it is merely a matter of this small difference in timing, it makes it look as if 'personal identity is likely not to be a nonderivative justifier' (p 460). But, says Johnston, the small difference in timing masks 'a crucial difference at the level of personal identity'. Reculturation is necessary before the Human Being can accept the Teletransporters' version of personal identity. According to *Simple Relativism* (pp 462-463) an individual is right to think of his personal identity surviving what he at the time believes that a person can survive while retaining the same personal identity. Johnston proposes *Modified Relativism* (p 463), according to which an individual understanding the possibility of future 'refiguration' is right to think of his personal identity surviving what 'refiguration' would enable it to survive. Johnston here seems to envisage a clash between what the individual believes (perhaps as a result of current social norms) and what he believes he will believe.

Johnston's (1989) arguments have the effect of loosening the common-sense concept of personal identity. He clearly regards 'refiguration' as possible. But he explicitly argues against those who use 'bizarre cases' to nullify 'the illusion of importance' created by the normal correlation of personal identity and psycho-physical continuities. Philosophers like Parfit (1984) may be represented as wishing to deny the validity of the concept[7]. Arguably what Johnston does is to

---

[7] See eg Cooper, 2008a, pp 205-206.

validate it by emphasizing its social elements, at the same time as suggesting the possibility of changing it.

Conclusions from this chapter

Writers who have paid attention to the social elements of personal identity show that these can be approached from at least three angles:

1. The empirical 'social roles' that James describes are elicited in response to different social groups. At times these healthy and useful variations can harden into pathological splits in personality, as considered by both James and Mead. Seen from this angle, there can be more than one personal identity associated with the same lifetime.

2. The 'spiritual self' or 'ego', considered as non-empirical by James, is represented by Mead as having a social dimension from the time it starts to develop. The concept of personal identity would not be the same without this social dimension in the self that generates it.

3. The concept of 'personal identity' is shared by members of the society speaking the language in which it is used. Johnston has suggested that the way this concept is used justifies a *relativistic* approach. His idea is that, as the understanding of personal identity varies from one culture to another, the person-directed concerns that go with it will vary as well. A Teletransporter will be concerned about what happens to the product of his teletransportation in exactly the same way as a young Human Being will be concerned about what happens to him as an old man. The *absolutist* approach, on the other hand, implies that there is only one correct interpretation of the concept. The 'absolutist Human Being', for example, would believe that a person cannot survive beyond his lifetime. He might be concerned about what happens to the product of his teletransportation but he would not think of this as concern about himself. On the *absolutist* view

the one correct interpretation of the concept must apply not just to all
contemporaneous societies but for all times. It would preclude the adaptation of
the concept to accommodate changing circumstances.

CHAPTER 3: HATCHING THE EGO

Introduction

In the introduction, I suggested a contrast between philosophers who emphasize the role of society in personal identity and those who emphasize the individual's role. Nozick is prominent among the latter, and I begin the chapter by considering his arguments. I go on to consider Parfit, whose analysis does not deny the social element of personal identity but tends to ignore it. The importance of Parfit's position for this paper is the contrast of his approach with that of Nozick and Cooper. Cooper accepts and elaborates Nozick's approach. However, he holds out a hand to the social analysis of personal identity. Still wading through the mud, I attempt to grasp it.

Two background questions on personal identity

For the discussion in this chapter it will be helpful to bear in mind two interrelated questions:

A) How does the self arise so that a concept of personal identity for that self is possible?

B) As asserted in chapter one, the concept of 'personal identity' is associated with that of 'self'. In what way are the two concepts associated?

The common-sense notion is that we do not know exactly when the self arises. We do not know how or when we become conscious of ourselves as persons, with the empirical 'constituents', material, social and spiritual, identified in James's classification discussed in chapter one. We can trace the evidence of stages in our physical development – birth, standing, walking, puberty and so on – with relative clarity. The stages in our psychological development are more obscure, though it seems reasonable to suppose they proceed gradually, like our physical

development. James identifies one such stage (1890, p 296), 'when we become able *to think of ourselves as thinkers*' (his italics). James says that 'at an early age' we can distinguish between a thought and what it is about. At this stage, when we start to think of ourselves as thinkers, we identify with the thought rather than its object. This must be a crucial stage for James, since he equates the 'spiritual self' either with the stream of consciousness or with the present segment of it, the Thought. An individual would have to have reached that stage to be able to identify himself as a person, that is, for personal identity to be applicable. But, if we accept Mead's account, discussed in chapter two, we can assume that the individual does not reach that stage in isolation from society. James's account does not preclude this interpretation, or deny the possibility that individual and social recognition of personal identity come at the same time. Nor does James's account belie the common-sense notion that self and personal identity develop gradually. James, e.g. in his chapter on the *Automaton theory* (pp 128-144), shows how consciousness is not epiphenomenal; it plays an important role in permitting more nuanced responses to the environment than nervous systems could achieve without it. It 'grows the more complex and intense the higher we rise in the animal kingdom' (p 138). For the human animal, this complexity and intensity is achieved slowly as the infant matures.

James, as I have just pointed out, identifies the 'spiritual self' either with the stream of consciousness as a whole or the present Thought. Cooper (2002) characterizes this theory as James's 'shoehorn'. The 'sense of selfhood', Cooper writes (p 121) 'is shoehorned into the immediately introspectible thought'. However, James is not wholly content to identify self with introspectible thought. He breaks his own rule about not venturing into metaphysics in a book on psychology (see James 1890, p vi) and develops the hypothesis of *sciousness* which I mentioned in chapter one. For Cooper, James's 'ultimate' view is that 'the minimal metaphysical self is *sciousness*, the passing Thought construed as pure-experiential fact' (p 125). James's words (p 304) are: 'The sciousness in

question would be the *Thinker*, and the existence of this thinker would be given to us rather as a *logical postulate* than as that direct inner perception of spiritual activity which we naturally believe ourselves to have' (my italics). As indicated in the introduction, there is a distinction to be made between logical postulates and metaphysical facts. To postulate something logically in order to explain an existing phenomenon is not to make an ontological claim. For James 'in an ulterior metaphysical inquiry' (p 304) there would be a question of '*who* the Thinker would be, or how many distinct Thinkers we ought to suppose in the universe'. So James himself contemplates the step from logical postulate to (potential) ontological claim. His contemplation is useful in showing the limitations of the postulate. We have to know who the persons are and how many there are in order to identify them, care for them etc. Such a postulate cannot tell us that, though it can help us to identify which spatio-temporal objects are persons or potential persons by facilitating our understanding of the concepts associated with it. I suggested in the Introduction that the existence of the self can be treated as a logical postulate. As such I think it helps us to understand concepts like 'self-consciousness'. But to talk about the 'existence' of the self in this way is not to make an ontological claim that there is a 'further fact' beyond the spatio-temporal object that is a person.

Background point on mysteries and beliefs
Nozick says a theory of the self should explain 'both the self's special awareness and its continuing mystery, even to itself' (1981, p 27). James also uses the word 'mystery'. He says (1890, p 296) that identification of the self with thought 'is a momentous and in some respects a rather mysterious operation'. The word has overtones which jar in a philosophical context; it is an instrument of tantalization. It suggests that some things are not known about the object or event to which it is applied, and hints that the reader would like to know what is not known but will not be told, or at least not until the end of the story. Nozick, at least, acknowledges that a philosophical explanation should remove the mystery. James, as Cooper points out (see chapter one) changes his position on the 'self' in

later writings. He may well feel that he is not telling the whole story in the *Principles of Psychology*.

Additionally, and more importantly, the word 'mystery' is imprecise. Something may be mysterious when we do not at present know it, although we could and perhaps will, know it. For example, we may come to know what is the cerebral activity associated with a particular thought. Other things may be mysterious because they are beyond the scope of knowledge, like Kant's *noumena* ('things-in-themselves'). Both types of mystery may apply to the self, but we need to be clear which type we are talking about at any one time. Neither Nozick nor James makes it clear. Indeed, James, occupying the dual role of psychologist and metaphysician, risks ambiguity as part of his job-description[8].

Cooper (2002, p 119) refers to Ayer's classification of beliefs in 'The Origins of Pragmatism' (1968), distinguishing those relating to logical or empirical facts from 'beliefs whose function is to satisfy our moral and emotional requirements'. James in 'The Will to Believe' (1896) observes that 'our non-intellectual nature does influence our convictions'. Such convictions – let us designate them 'non-evidential' - have what Nozick (1993) calls 'symbolic utility'. There is a social element to this kind of belief, clear in religious communities and cultures, where commonality of doctrine binds the group as well as distinguishing it from others. A problem arises when a strongly-held 'non-evidential' belief generates the hope that supporting evidence might somehow some time be found. In this situation, the distinction between the two types of 'mystery' can become blurred. A virtue of Kant is his emphasis on the limitations of human cognition, on what is beyond the scope of knowledge. He says, for example, '...if the pure concepts of the understanding try to go beyond objects of experience and be referred to *noumena*, they have no meaning whatever' (1783/1997, p 55). In examining 'further fact' theories (considered later in this chapter and in chapter four) it is important to

---

[8] Cooper (2002, p112) refers to Gerald Myers' 'surprise' at the apparent discrepancy between James's position on the self in 'The Principles of Psychology' and that of his later metaphysical writings.

keep clear the distinction between what is testable and what is not. Thomas Nagel (1974, p 441) points out that 'it is possible for a human being to believe that there are facts which humans never *will* possess the requisite concepts to represent or comprehend' (his emphasis). His context is 'realism' with regard to phenomenelogical experience. I would argue that his concerns apply equally to the field of 'self' and 'personal identity', and that one should be cautious even about using the word 'facts' for what he calls the 'humanly inaccessible'.

Nozick

Reflexive self-knowledge
Like James, Nozick associates the 'special awareness' of the self with self-awareness, and he thinks the self arises when a being can refer to himself 'reflexively', that is, use the word 'I' with full consciousness of the self to which the word 'I' refers. Nozick contrasts this with unconscious self-reference, as when Oedipus seeks the person who brought trouble to Thebes, without realizing that he is that person (Nozick 1981, p 72). Nozick says that 'reflexive' self-reference is 'reference from the inside' (p 75) and that this kind of reference 'bestows a property with the act of reference':

> It follows from its sense that the term "I" refers to the producer of that very token (of its type), and that person is referred to in virtue of the property he acquires in the very act of referring or producing the token, the property of being the producer of that token. It is part of the sense of the term "I" that it refers from the inside. (p 75)

Compare James (cited above) on the moment 'when we become able *to think of ourselves as thinkers*'. Both writers see as crucial to the self the particular form of self-awareness expressed in what Nozick calls reflexive self-reference. 'The linguistic facts' about how the capacity for reflexive self-reference is expressed, Nozick acknowledges (p 78), do not matter; but the capacity is 'part of the essence of selves'.

Let me elaborate a little Nozick's hypothesis for the way reflexive self-knowledge (the capacity for reflexive self-reference) comes about. He suggests (p 82) two possibilities for what reflexive self-knowledge is:

- It could be a 'basic phenomenon' or brute fact, incapable of explanation in terms of anything else. There would be no answer to the question of what quality enables selves to have self-knowledge merely ' in virtue of being identical with themselves' (p 82). Nozick discounts the 'brute fact' explanation. I do not, and I will come back to it.

- The second possibility is that reflexive self-knowledge could come about because the self 'places itself' (p 82) as the subject of predicates. That is to say the self could be a *non-linguistic* referring mechanism. We do not always just use words to refer. If I say 'That tree is beautiful', pointing to a tree, my reference is partly by means of the tree itself and partly by means of a gesture. The reference does not work unless I make the gesture and there is a tree there to which I can point. Indeed, I could make the reference just by the gesture, without using the words 'that tree', simply by pointing and saying 'Beautiful!'. Reference could be wholly non-linguistic. A predicate such as '... is beautiful' requires a subject to occupy the blank space and turn it into a meaningful sentence. Reflexive self-reference, Nozick hypothesizes, may be possible because the self could 'place itself into the blank' as 'part of a reflexively self-referring thought' (p 83). He seems to envisage a kind of mental pointing to oneself as 'reflexive self-reference'. If I am tired, I can say 'I am tired', or think 'I am tired' without saying it, and, when doing that, the self is 'mentally stepping forward into the space' in the predicate '...is tired'. However, Nozick acknowledges this explanation is circular. The self has to be reflexively self-aware already in order to know that it is itself that it is placing in the subject space. Self-knowledge is presupposed by self-reference and cannot be explained in terms of it.

A third explanation, favoured by Nozick, centres on 'the act of reflexive self-referring' (1981, p 87). As background to this explanation he introduces what he calls 'an abstract model of a procedure of classification' (p 84) used for grouping together 'things' (events, acts or physical objects). The basis for the grouping is degree of similarity represented in the model by the things' distance from each other in an 'n-dimensional space'. If there are n relevant ways in which the things differ from or resemble each other the space has to have n dimensions. Things are grouped together when they are close enough to each other and there is no thing outside the class closer to any one of them than they are to each other. In a 'maximally informative classification', distance between the members of a single group is minimized and distance between the members of one group and those of another is maximized. 'Entification' occurs when a group is identified. Nozick illustrates his model (p 85) with dots to represent 'things' and circles drawn round the dots to represent the classification (grouping). The circle 'represents entityhood' (p 84). According to one conception of classification offered by Nozick (p 85) an 'entity' is constructed in a moment ('time-slice space') according to the maximally informative classification at that moment, which is the 'closest relation' schema. To achieve identity of the 'entity' over time a 'closest relation' schema is again applied to 'classify stages as stages of the same entity'. Nozick, however, does not regard it as necessary to conceive 'entification' as taking place in stages. If we hold time to be one of the dimensions of 'abstract metric space', 'entification' can be seen as taking place 'in one fell swoop'. The distinction between ' time–slice' entification (which Nozick calls 'transverse identity') and 'entification' over time ('longitudinal identity') is then seen as an abstraction. But the distinction is 'natural' (p 86) because 'our entifications' are achieved from within time. Nozick says a maximally informative classification, by bringing together 'things' which are adequately similar, 'unifies', that is, creates a useful unity out of what would otherwise be diversity. 'Unity in diversity has been called organic unity'. Nozick emphasizes that 'entification', on his model, takes place within an assumed abstract space with dimensions representing degrees of similarity or dissimilarity of the 'things' being classified.

48

An explanation of the classification has to explain the dimensions of that abstract space, has to explain why the particular metric involved selects certain similarities and not others to be the basis for judgements about what 'things' should be grouped together.

This classification model is important for understanding Nozick's 'third explanation' of the emergence of self-knowledge. It is also important for understanding his 'closest continuer' theory, which I deal with a little later. For the present, I make two points about it. The first is that Nozick treats 'classification' and 'entification' as synonymous. This goes against the common-sense notion of classification. We are used to it as an aid to *id*entification of things that already exist but not as a means of creating something that was not there before. The second point is that classification is an abstract process, so, presumably, if it can create anything it can only create an abstract thing

I return to Nozick's 'third explanation' of how self-knowledge arises. He is concerned to identify the 'doer' of the act of self-referring and suggests the obvious candidate is the 'reflexive "I"' (what he calls 'the producer of that very token'). But he discards the notion of a 'pre-existing I' because this would lead to the well-known regression problem. (See the paragraph on Fichte below.) Nozick therefore supposes that the 'I' does not come into being until an 'entity is delineated and synthesized around the act A of reflexive self-referring' (p 87). The 'I' or self is that entity. It does not pre-exist reflexive self-reference but is created by it. A, according to Nozick, is an 'intentional action, embodied, (perhaps) in some physical production of sound or mark'. The act has 'dimensions' and 'components' around which 'a delineation of an entity' takes place (p88). The 'components' of the act include 'intention' and 'causal production' which provide the 'dimensions and metric' for the 'abstract space wherein the "I" is synthesized. A series of acts of reflexive self-referring, $A_1...A_n$, take place. They include the use of the 'closest relation schema 'bringing together things to constitute demarcated entities'. Another act A_0, which is 'partially' a

reflexively self-referring act, 'synthesizes' $A_1...A_n$ with itself, as 'parts or components or things arising from the same entity E' (p 89). E, the doer of all the acts, comes to exist in the synthesis of A_0.

'Synthesize' is one of the several words which Nozick seems to be using differently from common English usage without fully explaining what he means. Others are 'entify', 'delineate', 'dimension', and, indeed, 'act'. Here the 'act' is said to be 'intentional' but the agent, to whom any intention would be attributed, is created by the act. Cooper, to be discussed later, talks about 'an act without an agent' (2008a, p 141). The use of the word is suggestive of something more than an event but imprecise. I shall be arguing that part of what is being suggested is a 'further fact' beyond the physical facts that are being discussed. Nozick's description of his classification model does help the reader understand what he means. He seems to be suggesting that the acquisition of selfhood is like classification. Properties are grouped together (in the acts $A_1...A_n$) and simultaneously the groupings are grouped together (in the act A_0). This is the 'transverse' entification taking place in one 'time-slice'.

Nozick goes on to say (1981, P89) that by means of A_0 the self is 'positing itself as positing'. In the note on p 76 Nozick cites Fichte in support of this third explanation. He asserts that Fichte's claim that the 'I' posits itself 'includes at least' the notion of reflexive self-referring, which is referring to itself by means of 'a property bestowed or constituted' by the act itself. Fichte's suggestion of the 'I' 'positing itself as positing' is a denial that the property applies independently of the act. 'Can the self really be a Fichtetious object?', asks Nozick, without offering an answer. My answer is 'no'. Fichte does not envisage the self as a 'Fichtetious object' at all. He is aware of the regression problem. He notes that consciousness of the self as an object (in the sense of something perceived) requires consciousness of the self as a subject (perceiving that object). That requires the perceiving subject to become an object of consciousness to another subject 'and so on, *ad infinitum*' (Fichte 1797-8, p112). To solve this problem

Fichte suggests that there is an 'intuition' of 'an act of self-positing as positing' (p 113). This act is not just an act of positing; that would not overcome the circularity problem (because then another act would be required to posit what was doing the positing). It is not clear that Fichte solves the regression problem. (What enables the self to posit itself as positing?) It *is* clear that Fichte does not intend to suggest that the act creates an entity. He says (p 114) that self-positing 'is not supposed to produce an I that, so to speak, exists as a thing in itself and continues to exist independently of consciousness'. Rather, he regards this 'I' as a logical postulate, presupposed by self-consciousness.

Nozick suggests that some entity is 'created' by the synthesizing act. The language is admittedly ambiguous but my reading of it is that he takes what Parfit would call 'the Further Fact View' (1984, p 210). (Parfit himself says Nozick is a reductionist, holding that personal identity just consists in the 'closest continuer' relationship. Parfit's own reductionism is different, as we shall see below.) Parfit distinguishes between non-reductionists who believe personal identity is a 'Cartesian Pure Ego', and those who regard it as 'a further fact which does not just consist in physical and /or psychological continuity'. I will in the next chapter say more about the distinction between 'further facts' and 'logical postulates' in relation to the theory of personal identity.

Distrustful of Nozick's 'third explanation' of reflexive self-knowledge, I favour the first possibility that Nozick considers (1981, p 82) and rejects. This is the possibility that such self-knowledge is just a 'brute' fact. Consciousness always involves self-consciousness and this, in turn, implies the possibility of reflexive self-consciousness, the self's potential for expressing what it is doing or thinking. The realization of this potential must, of course, depend on the individual's ability to speak or express himself in some other way. But, according to this theory, it does not need some entity within the self, beyond the cerebral activity associated with consciousness. What remains 'mysterious' is the physiological mechanism which underlies consciousness of this kind. But facts about the correlation

between brain states and states of consciousness were already emerging in James's time and are emerging more rapidly with the imaging procedures now available to neuroscientists[9]. The 'mystery' is a lack of knowledge which we may eventually have. Belief in this theory is not non-evidential in the sense of being based on 'humanly inaccessible' facts.

Continuity

The continuity of the personal identity 'entity' is described by Nozick not as psychological or physical continuity but in terms of a relationship which he calls the relationship of 'closest continuer'. Nozick puts forward the 'closest continuer' theory as a general theory of identity and notes (p 34) that closeness is a necessary condition for identity, but whether it is a sufficient condition depends on the type of object being considered. 'The dimensions along which closeness is measured' also depend on the object. These dimensions, which are properties of the object, are weighted for the purposes of a judgement about identity. This does not mean, however, that all the properties which can contribute to closeness are necessary for a 'closest continuer' judgement. Nozick gives 'bodily continuity' as an example of what could be an important, even a determining, factor in some cases but is not a necessary condition for personal identity. A further element of Nozick's theory is that the 'closest continuer' relationship involves causal dependence. For x_2 at time t_2 to be the 'closest continuer' of x_1 at time t_1, x_2's properties must have developed from x_1's by means of a causal relationship. If, as you die, a random event produces an exact replica of you, this is not you (see p 41).

As Cooper points out (2008b, p 98), Nozick's theories of the 'closest continuer' and of the self as synthesized around acts of reflexive self-referring are integrated as an account of personal identity. Nozick imagines a series of synthesizing acts over time which 'mesh into a larger continuing entity' (1981, p89). Cooper likens

---

[9] See eg Barsalou et al. 2002

the continuing entification of the self, postulated by Nozick, to James's metaphor of a herdsman who inherits a 'legal title' to a herd of cattle and can identify the cattle by their brand (2008b, p 104). James's other metaphor, the chain that links the current Thought as 'hook' appropriating all preceding Thoughts (1890, p 340), also resembles Nozick's idea of a succession of 'acts' synthesizing and building on preceding acts. Yet James, at least in his discussion of the 'Pure Ego', tries to steer clear of metaphysics. (As I suggested in chapter one, I do not think the attempt is completely successful.) He claims that the stream of consciousness accounts for 'the unity, the identity, the individuality, and the immateriality that appear in the psychic life...as phenomenal and temporal facts exclusively' (p 344). The stream of consciousness is not 'the Soul', 'existing on a non-phenomenal plane' (p 345).

By means of those acts, according to Nozick's account, the self creates itself with the dimensions that it uses to recognize itself, to establish its own personal identity. 'Each person's own selection and weighting of dimensions enter into determining his own actual identity...I have special authority in fixing who I am' (p 106). Nozick's account, then, suggests that the self creates itself as an entity with authority over its own identity in deciding who its 'closest continuers' are.

Care

Nozick's view of care (Johnston's 'person-directed concerns') is related to his concept of 'closest continuer'. He says (1981, p 65) that we care especially about our 'closest continuer' when there is one, but do not particularly care whether there is one. Nozick takes it for granted that an individual cares about himself. He says that the self cares for the self with 'self-reflexive caring' which is caring 'in virtue of' being the synthesized self'. The synthesizing itself is 'an act of caring' (p 108). Nozick thinks this care will be projected on to the 'closest continuer'. 'We care about our closest continuer because we care about our continued identity, and that is what our continued identity comes to' (p 67). This association between 'care' and 'continuity' is reminiscent of James's claims about

the 'warmth' with which we feel for 'a present self and a self of yesterday' , as perceived in 'the stream of consciousness'. (See chapter one.) The idea is that the caring feelings and behaviour intrinsic to the concept of personal identity are *sui generis*. I would say that the individual who has developed personal identity cares both for himself and for others in this special way; it is part of the social interaction with which, as Mead shows (see chapter two) the self evolves. Common sense suggests that an infant cares from the moment it enters the world and starts bawling, but this is not the 'self-reflexive care' relevant to personal identity.

The question of care becomes more complicated when there are simultaneous continuers equally close, cases of 'ties', as Nozick calls them. Suppose an individual is superseded by two equally close persons, each close enough to be (the continuation of) that individual, if the other were not there. Nozick say we will care 'especially' about these because they 'best realize' the concept of continued personal identity for us (1984, p 67). Nozick distinguishes different progressive stages in the structuring of concepts (p 47 f). The 'closest continuer' concept is at the *closest relative* stage: it applies if nothing else is as closely related in the way implied by the concept. A further stage in the progression is *closest instantiated relation*: there is no other relationship which is as good or better as a realization of the conceived relationship. Nozick (p68) says that substituting the 'best instantiated realization' concept for the 'closest continuer' concept explains why we care especially for the closest continuer when it exists but do not care especially about having a closest continuer as long as we have adequately close continuers. 'I do not view a tie as like death; I am no longer there, yet it is a good enough realization of identity to capture my care which attaches to identity.'

Nozick supports his view that we do not care about getting a closest continuer when we do not have one by an analogy with our attitudes to contraception as opposed to infanticide (p 65). In many developed societies now contraception is

readily available and widely approved; infanticide is a criminal offence. Nozick's idea is that we do not care about potential children before they have been conceived and that, analogously, we would not care about potential 'closest continuers' until one 'closest continuer' had been produced. However, attitudes to contraception and infanticide have varied with different cultures and religions. Infanticide is still a matter of concern in some countries (see e.g. http://www.gendercide.org); contraception is forbidden according to Roman Catholic doctrines. 'Self-reflexive care' reflects social pressures and norms; this is an important way in which it differs from the egocentric care shown by the infant. Nozick spends a good deal of time speculating about how the individual's care will vary according to the 'distance' and number of continuers, but does not mention the impact on, or the input from, the society of which the individual is a member. The question whether someone has died or merely is 'no longer there' potentially has great medical and legal importance; Nozick writes as though it is up to the individual to decide.

Nozick also distinguishes two ways of caring (p 67). In the *Platonic mode* your care about a relationship is projected on to the best instantiated realization. In the second way you consider the best instantiated realization before deciding how much to care about it. Nozick, like Cooper, is a devotee of 'Platonic glasses'. Plato maintains that judgements about what occurs in this world depend on the existence of ideal 'Forms', such as the 'Form of the Good' and the 'Form of Beauty'. The Forms exist in some realm inaccessible to those in this world; what we perceive are pale imitations of perfection or as, the 'cave' image in *The Republic* suggests, shadows of the sunlit world beyond our reach. Through 'Platonic glasses', Nozick believes, 'we see the world in its aspect of realizing what is beyond it' (p 67). Also, we can see what is in the world as 'what it is the best instantiated (and a good enough) realization of' (p 633 note). Nozick acknowledges that the view through 'Platonic glasses' engenders 'an unrealistic overestimate of reality' (p 67). On the other hand, he warns that the second way of caring, looking first, makes you a 'a prisoner or victim of the actual world'.

It would be invidious to take this rhetorical passage as the last word on Nozick's *Weltanschauung*. But I will allow myself two comments on it:

1. 'Platonic glasses' increase the potentiality for confusion between evidential and non-evidential beliefs about which I worried at the beginning of this chapter. At the same time as encouraging you to accept what you have as the 'best instantiated example', they make you believe that there is something better, which you would ideally like to have and which you or someone else might eventually obtain.

2. The social perspective that this paper tries to adopt does not preclude the use of 'Platonic glasses'. Religious groups, for example, use them frequently, projecting on to real or imagined beings pre-existing concerns. However, it demands that such glasses should not be used all the time. The social interaction which launches and sustains personal identity is very much a practical matter. In so far as the practical pressures of the actual world imprison us we are all prisoners. It makes sense to understand the prison rules and see how the individual prisoner draws strength from the others while serving his sentence.

Parfit

Constitution of the self

Nozick's approach to personal identity is based on self-knowledge expressed in 'reflexive self-reference', 'reference from the inside' (1981, p 75). For him there is no such thing as personal identity or a person until the moment of reflexive self-reference. Parfit, by contrast, stands on the outside to consider how the self is constituted. 'A person's existence just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events' (Parfit 1984, p 211). This he calls the 'reductionist' view. Parfit also describes the view that there is a 'further fact'; personhood does not just consist in physical or psychological continuity. On this view a person is an 'entity *distinct*' from brain, body and events, in the same way as a nation is an entity distinct from its

citizens and territory. He is not implying that, according to this concept, a person could exist independently of the associated brain, body and events. He is suggesting that a complete description of reality could not refer simply to brains, bodies and events without mentioning 'persons' (see p 212). Parfit suggests that reductionists can accept the '*further-fact*' view and that the combination of 'reductionist' and 'further-fact' views comes closer than the pure reductionist view to 'our actual concept' of a person . What he means by 'entity' is not clear. He does not seem to use the word in the same way as Nozick. Perhaps Parfit's notion of 'entity' could be specified better in terms of concepts. Our concept of a person, he seems to be saying, is more than just a concept of the brain, body and associated events. The concept affects attitudes, so that it affects the psychological reality, which could not be fully described in terms of brain, body and events. The physical reality could be fully described in that way. I return to this point in chapter four.

Continuity

The difference between 'the further fact' as understood by Parfit and as understood by Nozick will need to be discussed later. I turn now to the question of continuity. On Nozick's view the continuity of personal identity is determined by the existence of one 'closest continuer'. For Parfit personal identity can be indeterminate in an 'unpuzzling' way (1984, p 213). He uses the model of a club which ceases to meet for a while and then is reconvened by some members with the same name and the same rules. It may or may not be considered the same club. The question whether it is the same club is 'empty'. Parfit says that we know all the facts without knowing the answer to that question.   Another way of putting the same point would be to say that the facts make the question inappropriate. Compare Parfit's club model with Nozick's story of the Vienna Circle (1981, p 32), an actual group of philosophers who left Austria when the Nazis came to power. Nozick supposes that three members of the Vienna Circle survive in Istanbul. On the assumption that all other members are dead, they can call themselves the Vienna Circle. But Nozick makes the additional assumption

that nine others escaped to the US. Had these continued to meet in Vienna it would have been they who were the Vienna Circle. Which group can now be called the 'Vienna Circle' depends on which group is the 'closest continuer'.

Parfit suggests that, like the identity of a club, our personal identity does not have to be determinate. 'Am I about to die?' can be an 'empty' question. When the reality is that body, brain and associated mental and physical effects are unified within one life, this view is counter-intuitive. But conceive a future in which processes like teletransportation, brain transplants and brain-contents transfers allow replications of personal identity, which 'overlap' in the sense that one personal identity is not destroyed immediately its replica is produced. If the question 'Am I about to die?' does not become empty, it does at least change in significance. We know all the facts and 'death' might not be an appropriate way to describe what happens. Again, compare Parfit's with Nozick's position. The latter (p 62) considers the case in which 'Y at t_2 and Z at t_2 continue X at t_1 equally closely'. Because there is no one 'closest continuer' he regards the most 'plausible' assessment as that X has ceased to exist. However, he does not regard a tie as 'like death'. (See the passage quoted above in paragraph 1 of my page 48.) [10]

Parfit, of course, is not saying that the absence of the 'closest continuer' relationship means the absence of all relationship. What he proposes is a relation 'R' which is 'psychological connectedness and/or continuity, with the right kind of cause, which could be any cause' (p 215)[11]. Note that Parfit distinguishes

---

[10] I have difficulty with Nozick's distinction between 'ceasing to exist' and 'death', as applied to personal identity. Cooper's explanation may help. 'The integrated Closest-Continuer theory does *not* imply that hypothetical cases of cloning or branching, familiar in the literature of personal identity, are tantamount to death when the branches are equally close to the originating or «trunk» person' (Cooper 2008b, p 100). Cooper argues that through 'Platonic glasses' (Nozick's phrase) one can see beyond the actual world and could see branching as survival in order to avoid oblivion.

[11] This is another point of disagreement between Nozick and Parfit. Suppose a replication machine has a whole stock of all the psycho-physical data required to re-create personal identities. Every time someone dies it selects at random one set of all the different categories required for one personal identity. When you die it happens to select the set that replicates you. Nozick would not

between 'psychological connectedness', such as having direct memories , and psychological continuity which is having 'overlapping chains' of connectedness. This psychological continuity is 'what matters' (p 262). Physical continuity is relatively unimportant, though Parfit suggests (p 285) that the 'R' relation requires sufficient physical similarity to allow 'full psychological connectedness'. A brain or brain contents transplanted to a completely different body might have difficulty in functioning as before, but how far 'psychological connectedness' can be damaged in this way before the relationship ceases to hold is unclear, perhaps impossible to prescribe.

The strength of 'psychological connectedness' does not just vary in cases of replication. In an earlier paper *Personal Identity* (1971), Parfit imagines 'immortal beings' for whom direct psychological connectedness would hold only for limited, overlapping periods. These, he suggests, would not use the word 'I' to refer to long-past experiences but would speak of 'earlier selves' (p 25). We could, he says, think about our own lives in this 'surprisingly natural' way. Parfit talks about 'successive selves'. David Lewis (1976) suggests the use of the phrase 'person stage', the stages being linked by a relation 'I' comparable to Parfit's 'R' relation. Also – again, comparably to Parfit's 'R' relation - the strength of the relationship varies. For example, the stage several hundred years through the life of Methuselah would be weakly related to the first stage. Hence 'personhood: the property of being a continuant person' is a matter of degree (Lewis, 1976 p32).

More important, from the point of view of this paper, is the extent to which the 'R' relation can deviate from personal identity. Parfit says

     PI = R + U

where PI is personal identity, R is the relation, and U is uniqueness, 'the fact of a relation holding in a one-one form' (p 263). It is the relation, Parfit claims, which produces 'most of the value'. The presence or absence of U does not make all that much difference. He thinks that the case of 'division' proves his point.

_____

count this as your closest continuer (see 1981, p41, quoted above); Parfit would.

Continuation as two persons would matter to the individual even though his personal identity had ceased as a result of the subtraction of U. Parfit seems to be suggesting that we should play down the emphasis on personal identity in order to concentrate just on the relationship. An alternative would be to adapt[12] our concept of personal identity to emphasize the relationship element rather than the uniqueness element. Parfit looks at the concept and the relationship R very much from the point of view of the individual. In the next chapter I shall discuss how the concept could be adapted to supply social needs.

Care

I have argued that what Johnston calls 'person-directed concerns' and what Parfit calls 'special concern' (p 307f) are an important aspect of the personal identity concept. If this is so then it would seem that Parfit might accept adaptation of the personal identity concept rather than playing it down. In any case he rejects what he calls the 'extreme claim', that a reductionist does not have reason to be concerned about his own future, in favour of the 'moderate claim' that a reductionist's concern is centred on the relation R. This accords with the common-sense view that psychological continuity will always be associated with concern for the future just as it will be associated with concern about the past. If we remember with pleasure or pain the past physical and mental events associated with our brains, we will surely care about the effects of future events. Such care does not seem necessarily to depend on the uniqueness of personal identity as we know it. Parfit suggests that caring about the future of multiple personal identities might be compared to caring about loved ones (p 312), but is not the same as anticipating e.g. pain we will suffer ourselves. The depth of care would perhaps depend on the number and propinquity of the personal identities. Having too many could generate hostility; having one or two available close at hand could be convenient and reassuring. The existence of multiple personal identities would

---

[12] Concepts change as circumstances vary and knowledge increases. Brigandt (2009) argues that a concept has an 'epistemic goal' according to which it is rational for the concept's reference and inferential role to adapt.

also affect the concerns, both emotional and practical, of the society around the individual. The adaptation of the concept would have to take account of these effects.

Cooper

Origin of the self
Where Nozick summons up the ghost of Fichte, Cooper invokes Descartes. The much-debated *cogito –*'I think*, therefore I am' -* was criticised by Lichtenberg, who regarded Descartes as only entitled to say 'it thinks' or 'there is thinking going on' (*Aphorisms* trans. 1990). Cooper (2008b, p 102) suggests that *cogito* implies '*This very thought'* is going on. The thought is aware of itself and this act of (what Nozick calls) 'reflexive self-awareness' *bootstraps* itself into a subject of experience, a self is *entified*.

This Nozickean account of the self's origin is elaborated and refined in Cooper's *An Eldritch Tale: Kafka's The Metamorphosis and the Self* (2008a). Here the 'act without an agent that bootstraps the self into existence' is personalized with a name: Bootstrap. Bootstrap is 'internal to consciousness' (p 141). There is no 'anterior causation' but this does not make it anomalous because 'it remains possible that there is a neural correlate of the self's agency, simultaneous with its exercise, and that the correlation is implied by a scientific law.'   Cooper does not elaborate on this possibility here, but, in his work on James (2002, p 74), he uses Alvin Goldman's concept of 'simultaneous nomic equivalents' to explain James's approach to the relationship between mental and physical events. The suggestion is that a particular 'bodily process' is the 'nomic equivalent' of a particular mental process ('stage in the stream of consciousness'). There is a 'scientific law', according to which the one cannot occur without the other. '...mental principles and brain laws are co-determinants of behaviour' (p 78). This account, in Cooper's view (p 79), implies 'laws of correspondence' (dictating that things must happen together) but not 'determinism' (dictating that one thing must be

followed by another). Determinism would suggest there has to be an antecedent cause.

This use of the 'nomic equivalents' theory, in my view, smacks of overdetermination. If all the firing-squad bullets are forced by a scientific law to hit the same spot, you do not need a firing squad. Saying that mental and physical principles are 'co-determinants' that have to act together is not very different from saying that they just represent two ways of conceiving the same principles, as a physicalist might. Cooper however is specific in rejecting physicalism. He denies Gerald Myers' reductionist interpretation of James. 'Correspondence is not identity' (Cooper 2002, p 85). In *Nozick, Parfit and Platonic Glasses* (2008b, p 102) he calls the Nozickean self-defining entity 'the Real Self' and says this cannot be reduced to a physical description. Parfit, as indicated above, considers a reductionist view of the self is compatible with a 'further fact' view. Cooper does not.

Another point about the 'simultaneous nomic equivalents' theory is that it does not overcome arguments about the 'anterior causation' of mental events. If a particular mental event has to be accompanied by a particular physical event, and that physical event is subject to the laws of physics, then the mental event is also subject to those laws. Cooper acknowledges that the theory is compatible with the notion of 'the body as a reflex mechanism' (p 79), but says that James thinks the brain is different from the rest of the body in this respect. Cooper must agree with James here to sustain the idea of no 'anterior causation'.

Uncertainty about the origin of the self is enhanced by Cooper's entertainment of alternative accounts. The 'Bootstrap' account I have just described he terms 'full reduction' (2008a, p 143). Here he is not using 'reduction' in Parfit's materialist sense but, rather, in the sense of excluding from the self anything but the act without an agent. The alternative to this is what Cooper terms 'partial reduction', 'in which the self emerges as entified in the psychological development that

creates the sense of self.' The emergent self is personified as 'Starter' in Cooper's *Eldritch Tale*, who 'causally emerges from ontogenetic development in the brain' and 'may be conceived as a momentary self' or as 'Bootstrap', the act without an agent (p 140). The 'partial reduction' approach, at least, allows for the idea of gradual development, which fits better with the notion that social forces impact on the way the self emerges.

Continuity

Both Nozick and Cooper visualize the self as episodic rather than continuous, in the way that James envisages the 'stream of consciousness' as a series of individual 'thoughts' (or episodes of consciousness). In *Nozick, Parfit and Platonic Glasses*, Cooper describes 'the Real Self' as following the pattern of 'episodes of reflexive self awareness'. 'Entifying these episodes into a continuing person is a step that is justified by viewing the episodes through Platonic glasses' (2008b, p 104). Platonic glasses or not, the notion that the self is somehow in abeyance when the individual is not aware of himself is counter-intuitive. A physical system is not, of course, subject to intervals of non-existence. So a reductionist does not need 'Platonic glasses' to justify his belief that he continues as a person while asleep or that what happens to him during intervals of unconsciousness matters.

In *An Eldritch Tale* (2008a) Cooper uses for his model Kafka's story 'The Metamorphosis', in which Gregor, a previously normal young man, wakes to find he has psychological continuity but has been physically transformed into a creature with the body, movements and appetite of an insect. Cooper calls the self at the time of the metamorphosis 'Inbetween'. 'Inbetween' has to decide whether to remain Gregor or become a different person, 'Beetle'. 'Inbetween' and 'Beetle' seem to belong to a 'partial reduction' account. The entified self makes the decision about which personal identity to adopt. Cooper says (p 135) that it would be 'rationally permissible' for Inbetween to decide that Gregor had ended and Beetle had begun if the metamorphosed creature were unacceptable as a

continuer of Gregor. 'The rational permissibility in question is what I shall call the rationality of individual authenticity, in order to distinguish it from the rationality of social approbation.' The distinction is between what changes can be tolerated by an individual before he claims he is a different person and what changes are tolerated according to the standards of society. In the case of the 'Teletransporters' envisaged by Johnston an exact replica could not be accepted as a continuation of personal identity unless the community recognized the method of replication used. Cooper notes that there *may* be cases in which 'the individualistic and communitarian forms of relativism about the self converge', as if this might be accidental or rare. Convergence, however, is the norm for the very good reason that 'individual authenticity' has its basis partly in 'social approbation'. 'Individual authenticity' comes from the 'metric' that Nozick discusses as well as 'personal concerns'. Both metric and concerns grow out of the individual's responses to the people around him and their responses to him. This theme is expanded in the next chapter in the sections on *The emergence of the self* and *Personal concerns*.

An 'episodic' account of the self, like Nozick's or Cooper's, may seem to require an explanation of how the episodes are joined together[13], that is, how continuity is achieved. Nozick has the idea of the 'closest continuer' relationship such that each person has 'special authority' (1981, p 106, quoted above) in deciding that the current episode is his 'closest continuer'. Cooper identifies three 'forces that bind the self together through time' (2008a, p137). In addition to the 'metric' informed by 'symbolic utility' (basically Nozick's idea of how the 'closest continuer' is selected), for Cooper there are 'habit' and 'expected utility' to be considered. 'Expected utility' is surely the forward element of the psychological continuity of which memory is the backward element. To be conscious involves planning the future as well as remembering – and making use of - the past. As for 'habit', Cooper is using the word , as James does, for the conditioning of the

---

[13] Another view is that it does not require such an explanation, any more than an explanation is required of how drops of water are held together in a stream. Drops and stream are just two different ways of looking at the same phenomenon.

nervous system brought about by repeated exposure to the same stimulus. Some such 'habits' derive from innate tendencies, others from training (see James 1890, p 104). The latter is, of course, deliberate and, for the child especially, very much a product of social exposure. James himself calls 'habit' the 'fly-wheel of society' (p 121). James's account of 'habit' is also 'physicalist': 'nothing but *concatenated* discharges in the nerve-centres, due to the presence there of systems of reflex paths, so organized as to wake each other up successively' (p 108). I might say that the Nozickean 'authority' of the individual is limited by his psycho-somatic constitution as well as by social constraints; Nozick might say it is enhanced by his psycho-somatic potential and ability to ride with and above social forces. Anyway, there are internal and external pressures to remain consistent and, normally, internal and external boundaries to the changes possible for an individual. Cooper himself acknowledges 'limits to the expansive subjective range for self-definition' (2008a, p 138).

Care

In *Nozick, Parfit and Platonic Glasses*, Cooper shows that for him, as for Nozick, care is bound up with the 'closest continuer' theory, and with the use of 'Platonic glasses'. 'Symbolic utility' (which I would say implies and is implied by care) plays a part in binding the episodes of self together in a 'normal' life. In the abnormal situation in which there are 'branching closest continuers' (two or more equally close continuers) the 'rational' response is for the individual to care about the branches as continuations of himself (2008b, p 101). Care, for Cooper, might take the form of bequeathing equal portions of one's wealth to each of the branched 'clones' or disinheriting all of them in favour of a charity. (See 2008b, p 103). This scenario demands a social imprint. There would have to be laws about how wealth is taken over by replicas. *An Eldritch Tale* (2008a) confirms and elaborates this approach to care. Cooper says that continuation in accordance with the three 'binding' factors (habit, the metric and expected utility) is 'close enough to care about' (p 137). It is both 'intrinsically desirable' and 'instrumentally valuable' for achieving goals. The distinction is important, not

just for the individual's attitude to his own personal identity but for society's as well. For families and communities personal identities have 'expected utility' with regard to the demands that we make on individuals and to enable us to hold them to account. We also accord intrinsic value to them, respecting and caring for individuals as 'ends in themselves', to use Kant's phrase.

Conclusions from this chapter

The Nozick-Cooper 'bootstrap' take on the origin of the self may be seen as inimical to the idea of social input. According to this view (put crudely) the self seems to create itself out of nothing in an instant and in isolation. It then continues on its own authority deciding at each stage whether it is the same self or not and, if not, what self it has become. Cooper acknowledges Johnston and the socially relativist ideas about personal identity, and he allows that the two sorts of account may 'converge'. But we need a single account which gives adequate emphasis to social as well as individual elements of the self's origin.

It is no doubt clear from what I have already said that I prefer accounts which emphasize the gradual development of the self. With these it is easier to point out the impact of social factors at every stage. Accept Parfit's account, according to which the individual's personal identity is basically his individual brain and body, and you also have to accept the associated mental and physical events as part of the package. These events consistently involve the society around the individual who is reacting all the time with or against his social environment. As I have suggested before, for a brain and body abandoned at birth in an inanimate environment there would be no concept of personal identity as we know it.

Possibly, however, this picture of the deserted babe can provide clues to ways of socializing Bootstrap. I will enlist help, too, from the block of marble arranged in the shape of a man on horseback (see footnote 7), which acquires its significance from the interaction between the sculptor's intention and the people who see it. Even an act without an agent is not an act without a context. If such an act can

take place, so to speak, inside a human being's abstract space, it takes place in relation to a being who has evolved to be a social being, whether he likes it or not. Even if Bootstrap becomes reflexively self-aware by himself he must at the same time become aware of his circumstances, particularly the people around him, that give significance and value to the self-awareness. The metric that helps to shape his continuity is influenced as surely by his social circumstances as by his own body.

CHAPTER 4: PANORAMA WITHOUT PLATONIC GLASSES

Introduction

This chapter aims to:

- elaborate and clarify two themes: the emergence of the self and personal concerns
- identify conclusions from the discussion as a whole.

The emergence of the self

In this section I will mainly consider empirical features of the emergence of the self, that is features which can be experienced either through perception or through introspection. Towards the end of the section I will consider some of the metaphysical arguments about the basic factors underlying these experiences.

I see this emergence as taking place over a long time, during which the properties of the self are developed in a social context. It is these properties that James associates with the 'abstract empirical self' mentioned in chapter one. The phrase refers to the 'psychic faculties or dispositions' (James 1890, p 296) which are essential to an individual's character. His awareness of them is, of course, immediate. He does not have to study himself to know that he is conscientious or ill-tempered. Even the knowledge that something is a long-term disposition, as opposed to a transient feeling, is built up automatically, in a sort of conditioning process, and not as a result of deliberate research. The tendency to behave consistently, 'habit', is a component of the 'glue' holding personal identity together and discussed in chapter three. As James points out, it is often unconscious, thereby saving too much conscious thought. But it comes into consciousness, as when it clashes with the individual's conscious goals or those of

society. James, as I indicated in chapter three, sees consciousness as permitting more complex responses to the environment from the nervous system. Nozick, too, stresses the role of consciousness in his discussion of reflexive self-awareness. For a great deal of time, from the beginning of consciousness, conscious behavior takes place in a social context. Normally humans are consistently in company, and they acquire knowledge of themselves at the same time as they are acquiring knowledge of others and others are acquiring knowledge of them. As the process continues the individual is immediately aware of his reactions to other people and mediately aware of their reactions to him. James's emphasis is on self-consciousness as a 'reflective' process. My emphasis is on its interactive nature.

Mead asserts that selfhood is achieved by the individual's becoming 'an object to himself', 'taking the attitudes of other individuals towards himself' (1934, p 137). For Mead the self emerges automatically as the child attains, in interaction with those around him, the self-knowledge which enables him to become a person with a personal identity. Not all the beliefs that someone has about himself constitute knowledge; they can be mistaken, in spite of the immediacy of the sources of the beliefs. There is a sense in which self-knowledge is infallible, as, for example, when I am aware of pain. If I claim sincerely that I am in pain, the claim is justified, whether or not there is an ascertainable source of pain. But there are ways in which self-knowledge is fallible. I can have false memories. I can become deluded, as when – an idea discussed in chapter one – I think I am Napoleon. In such circumstances the interaction between the individual and society still holds but goes into a different gear. With their more objective knowledge of me, those around me might correct my false memories, might persuade me that I am not Napoleon. If their persuasion is unsuccessful, they will have me locked up. But suppose I 'really' am Napoleon, in the sense that Napoleon's preserved brain has been implanted in what was EH's body. In that case, I will be like Locke's cobbler. I will have some fast talking to do or I will still be locked up.

69

Mead explicitly differs from James in his understanding of how the self emerges. The former says, as I pointed out in chapter two, that James regards 'reflexive affective experiences' as the basis of the self (1934, p 173). Mead down-grades the importance of these experiences. 'The individual need not take the attitudes of others towards himself in these experiences, since these experiences merely in themselves do not necessitate his doing so, and, unless he does so, he cannot develop a self.' However, Mead's comment needs qualification. The 'self-feelings' or 'self-estimations' of which James gives examples in his table (see chapter one) are pride and modesty. Both feelings are influenced by comparison with other people and exposure to praise or blame. Someone growing up in an uninhabited world could feel neither pride nor modesty of the usual kind. He would feel pleasure in the ability to get what he wanted and dissatisfaction when he could not get it, and these could be long-term 'self-feelings' or dispositions as opposed to short-term affective reactions. But they would be less complicated than pride and modesty, the lack of complexity coming from lack of exposure to other people.

Anyway, although it is true that James emphasizes the reflexive nature of the self in his account, for instance, of the stream of consciousness, he does not disregard the social aspects of the self and, as I have suggested before, he would not have rejected the idea that social influences are significant. His discussion of the 'social' component of the 'empirical life of self', dealt with in chapter one, shows the importance he attaches to social roles and the way the individual performs differently in different social arenas. His discussion of the 'spiritual' component of the 'empirical life of the self' (1980, p 296) shows the importance he attaches to that. Without it a man becomes *alienatus a se*. Without it, one might add, a man also becomes alienated from the society around him. The dispositions identified by James's 'abstract' approach to this component are accessible to society and amenable to the influence of social interaction. James's 'concrete' approach to the 'spiritual' self identifies the celebrated 'stream of consciousness'.

A human stream of consciousness is shaped by the conscious human environment, not just the self but the external social world. Mead notes (1934, p 138) that in the development of reason the individual comes 'into the same experiential field as that of the other individual selves in relation to whom he acts in any given social situation'. It is through this 'field' that the stream of consciousness flows.

The theories discussed in this paper tackle not just the empirical facts of the self's emergence but also the metaphysical question of what lies beneath these facts. I will consider some points about the metaphysical discussion which are relevant to my thesis.

Parfit's account (1984), described in chapter three, reduces a self to a 'brain and body' and a 'series of interrelated physical and mental events'. Add to this the point that the mental events *are* physical events. A thought could be completely described in physical terms, given sufficient knowledge of the neuro-physical activity associated with it. To use everyday language and to call this activity 'my thought' is to apply a mental concept to the physical activity of my body and, by implication, bring into play other mental concepts. Thoughts are associated with other mental events and with dispositions and so on. But to use these terms is just to apply mental concepts to physical facts. There are no 'further facts'.

This physical account of the self is compatible with my thesis. From the moment they emerge from the womb, brain and body are nourished by the social environment within which personal identity develops and continues. As I said in chapter two (as a proffered resolution to the supposed Mead-Nozick conflict) the individual achieves personal identity by becoming *both* an object and a subject to himself. This duality – being object and subject – is just a fact of self-consciousness and self-consciousness is just a fact of consciousness. As Fichte points out, 'I can be conscious of any object only on the condition that I am also conscious of myself, that is, of the conscious subject' (1797/98 – 1994, p112). An analogy might be illumination, in which the source of light illuminates itself as

well any objects exposed to it. There are problems in describing consciousness.
Language has evolved to describe that of which we are conscious, rather than
consciousness itself. There are problems in understanding consciousness, not
least in understanding how it develops to the point at which a conscious individual
can be said to have personal identity. But the problems themselves do not mean a
physical account cannot be adequate. On such an account, the gradual
development of the self would just be the gradual development of the brain and
body. Both grow healthily in interaction with the social, as with the physical,
environment. Cut off the social environment and distortion occurs.

Many philosophers of mind, from Plato to the present, have been dissatisfied with
a purely physical account of self. Reasons for this dissatisfaction include:
A. the idea that humans are superior creatures, designed by God to be different
from other animals
B. the idea that human experience cannot be reduced to physical terms.

Philosophers of the Nozick-Cooper school seek 'further facts' to explain personal
identities. Nozick claims that 'selves are special among entities in having their
identity over time be (partially) self determined' giving everyone 'a special, and
desirable, dignity' (1981, p 107). The point of my argument here is to suggest
that the Nozick-Cooper version of the 'further fact' account of personal identity is
not just a restatement of a physical account in a different sort of language, but an
attempt to provide something more. Parfit's version of a 'further fact' approach,
discussed in the previous chapter, *is* a restatement of the physical facts, albeit a
restatement with affective implications. A nation feels different from a random
group of people deposited on a random bit of territory, and inspires different
feelings in others. That does not mean that there is anything more than the people
and the territory and a series of events relating them. A complete description of
the people, territory and events, with sufficient psychological and historical
insight, would show the difference between the nation and the random group and
explain the affective differences.

The Nozick-Cooper type of 'further fact' is not independent of physical events. These philosophers do not suppose, as Descartes supposes, that the self emerges because the human being has a soul, consisting of a separate substance and able to exist independently of his body. Yet, they represent the individual as creating his own personal identity or having it created for him (the 'act without an agent') in the moment of reflexive self-awareness or metamorphosis. They appear to envisage an individual creative act in addition to the socially interactive psycho-physical development I have been suggesting. What I have called 'an addition' could just be a difference of emphasis. An analogy might be an account of the 'birth' of a nation, say Israel, as the creation of a few inspired Zionists, dreaming of the 'promised' land. Such an account would exclude many of the crucial historical, geographical and demographic factors from which Israel arose. Yet the inspired Zionists were there. On my interpretation, however, the Nozick-Cooper approach is not just a difference of emphasis.

Assuming that the 'Nozick-Cooper further fact' is an additional fact, one possibility is that it is among facts which we have not yet accessed. Cooper's 'neural correlate of the self's agency', which lacks antecedent cause but is not anomalous, is a hypothesis open to proof or disproof. Proved, it would rank as a fact. Souls rank as what Kant calls *noumena*, things in which we can believe for good or bad reasons but not facts about which we could know. To the evidential 'further-fact beliefs', those that could in theory be proved or disproved, what I said about the compatibility of my thesis to the physical account of the self applies. With or without antecedent causes, the 'neural correlate' would have to function within a human body exposed to the social environment.

The alternative possibility is that the 'Nozick-Cooper further fact' is among the *noumena*, something we cannot know about in the way we can know about physical facts. If so, belief in it must be justified by its 'symbolic utility' (see chapter three).

Person-directed concerns

I said at the beginning of the paper that two aspects of the self relating to personal identity are 'continuity' and 'care'. I have said a lot about 'continuity' of personal identity, comparing different views of what constitutes continuity and how it can be achieved. I shall, in this section, say more about 'care' or what I have tended to call 'person-directed concerns', using Johnston's phrase. Let me recapitulate Johnston's list of these concerns and expectations (1989, p 448), quoted in chapter two:

> '(i) one's future oriented and retrospective concerns for oneself and others;
>
> (ii) one's expectations about experiences and memories of those experiences;
>
> (iii) one's expectations about the relations between action and desert.'

The list will serve as a reasonable guide to the sort of concern we are talking about.

It also serves to show how the individual's person-directed concerns are mirrored in society's concerns about him. This is obvious in group (i): the individual has 'future oriented and retrospective concerns' for others; others have them for him. As far as group (ii) is concerned, a large proportion of the normal individual's experiences take place in a social context, so that the individual's memories and expectations are allied to those of the people around him. To cite Mead again, the individual enters 'the same experiential field as that of the other individual selves in relation to whom he acts in any given social situation'. The individuals' memories of their own history and plans for their own future are, of course, individual. But they are interdependent. When old friends come together to reminisce, this is not just idle chatter but, at a deeper level, mutual reassurance about the continuity of personal and social identity. The individual's plans, for anything from a holiday to a career, depend on the other people involved doing what is expected of them. It is in such ways that memories and expectations are allied. If the alliance is betrayed, as with Dr Jekyll or Gregor Samsa, big

74

problems ensue both for the individual and society. When someone is stricken by a disease, such as Alzheimer's, which destroys his ability to remember and to plan, it is not just the individual who is affected.

The expectations in group (iii) are moral beliefs, also shared with society. Indeed, the individual acquires these beliefs in the process of his moral education. He may come to prefer his own principles, as James points out, when they clash with other people's. But the social roots of moral beliefs are not in doubt; the individual in an isolated environment may have expectations about the effectiveness of his actions but not about his 'desert' as something distinct from his 'desire'. Such moral concepts presuppose the existence of other people or, at least, of other beings.

To be a person at all involves having person-directed concerns. The newly-born infant has needs but not concerns. The latter develop as it learns how to satisfy its needs in the social environment in which it finds itself. (It might be said that the infant abandoned in an uninhabited environment would, if it survived somehow, develop its own 'concerns' but these would be formed and displayed differently. Mead, as noted in chapter two, allows that such an isolated individual might represent 'a certain feeling-consciousness', but it would not be a normal self.) Whether the self emerges gradually, as I have argued, or suddenly, as Nozick and Cooper suggest, the concerns are a prominent feature of the personal identity and of the 'metric' that, as Nozick and Cooper both maintain, the individual is at pains to preserve. Important components of a metric must be the 'dispositions' that James talks about and personal concerns. The two categories overlap. X's concern for the detailed characteristics of the trees goes hand-in-hand with his pedantry over minutiae; Y's concern for the overall layout of the wood is inseparable from his penchant for sweeping generalizations and impatience with detail.

Much of my discussion of personal identity has been about what constitutes continuity for a person. Most of the time I have taken it for granted that a person wants to continue. Such an assumption seems justifiable, with obvious exceptions such as someone suffering from a painful and incurable disease. But for persons there is more to the desire for continuity than the instinctive wish to remain in existence. We have the deeply held conviction that that about which we are personally concerned is worth preserving. Such convictions can trump the animal instinct for physical self-preservation. People sacrifice their lives for others or for political or religious principles. If non-evidential beliefs about how personal identity is preserved are justified, it is because they serve these convictions. Belief in the immortality of the soul, for example, may help those who have it to pursue their concerns without worldly rewards. If belief in the 'Nozick-Cooper further fact' is non-evidential, it must be justified in some such way. Parfit regards Nozick's 'closest continuer' theory as reductionist and denies that Nozick believes in a 'further fact'. (See *Reasons and Persons*, Appendix E, 1984, p 477f.) Nevertheless, he questions the rationality of Nozick's view that unique identity, as depicted in the closest continuer theory, is what matters. Parfit holds that it is the relationship 'R', rather than uniqueness, that matters (see chapter three). However, he concedes that if Nozick's belief is 'more deeply satisfying, it can be practically rational for [Nozick] to try to make himself, in this way, theoretically irrational' (p 479). I do not accept that irrationality can be rational, but non-evidential beliefs can be justified, in the way that Parfit justifies Nozick's belief, in virtue of their effects. For Johnston's 'Human Being', discussed in chapter two, the availability of 'teletransportation' is not enough to secure his belief in continuation unless he undergoes 'reculturation' by the 'Teletransporters'. Johnston's 'modified relativism' offers a kind of limbo in which the individual can hope for survival without being sure of it. Whether you are Nozick or Parfit, this position has symbolic utility in so far as you do not like the idea of your personal identity ending with death. Personal concerns such as the wish for continuity have expected as well as symbolic utility.

I have said that person-directed concerns mirror the concerns of society. They do this in obvious ways. James's empirical 'social roles' are played out in response to society's demands and society responds to them. The individual is concerned about whether he will be rewarded or punished; the past actions that he regrets or in which he takes pride, and his plans for the future, usually correspond to those actions which society seeks to punish or reward. Society needs to be able to reward and punish individuals. These actions have both symbolic utility as expressions of approval or disapproval and expected utility in encouraging or discouraging types of act. It follows that the same concept of personal identity has to cater for the concerns of both the individual and society. If the individual does not have the same concept as society it is problematic. For Mary Reynold, the case about which James wrote and which is discussed in chapter one, it is not enough that she has two personal identities from her *own* point of view, as it were; unless society accepts there can be two personal identities in the same body, she will continue to be treated as one abnormal person.

There is also a less obvious relationship between society's concerns and the individual's. The self-interest that motivates the individual to seek the continuation of his metric is also the interest of human society, which has evolved to require continual change and therefore continual individual differences. I am not going to speculate about whether this principle applies to other forms of life. But it applies to human animals that a wide pool of different ways of dealing with the problems of survival is an essential resource. In this sense, therefore, person-directed concerns are society's concerns.

It follows from all this that a society offering 'refiguration' has to have a coherent system of beliefs to support the refigured identity. It is not only the case that the individual must cease to regret what he did but can no longer remember (because it was in a distant 'cryptophase' of his continuation, to use Johnston's term). Limited memory already achieves something like that. Additionally, society must forgive the individual for what he did in a distant cryptophase. The alternative is

surely as intolerable as the belief that the sins of the parents should be visited upon the heads of the children. I suggested in chapter two that Johnston's arguments 'loosen' the concept of personal identity. In effect, much of the current philosophical discussion of the concept does this. My point is that the loosening should be extended to social aspects of the concept which will need to be accommodated in any settlement of the questions about personal identity.

Conclusions

My aim in this final section is to draw together some of the main arguments from this and the preceding chapters in order to summarize the implications of the 'social' approach to personal identity.

Metaphysics

I have not attempted to settle the question of what the 'self' is; I have merely tried to distinguish 'reductionist' and 'further fact' accounts and to suggest that the 'logical postulate' approach to explaining the self is not a metaphysical account. From some points of view, the absence of an agreed metaphysical account does not matter. The concepts of self and personal identity can both be used and developed irrespectively of the metaphysicians' debate. This is how 'everyday' concepts work. We are able to drive automobiles without a deep understanding of internal combustion engines. We are able to talk about selves and personal identity and use ideas like continuity and concern without understanding their metaphysical basis.

My preference is for a 'physicalist' account of the self. If it is correct, the basis for the concept of personal identity is, in theory, capable of empirical investigation, which may eventually be possible in practice. We can imagine a brain scan, say, vastly more sophisticated than anything available at present,

which would provide a detailed record of an individual's brain state, highlighting the elements underlying personal identity. We can also imagine that long-term investigation has brought agreement about which elements these are. This scan would provide a far more accurate representation than a qualitative description given either by the individual of himself or by the people around him or both. Suppose in this possible future – the year 2984, perhaps – replication is readily available but sometimes only partially successful. The replica in which someone has invested, prior to assisted suicide (which is also readily available), may not reproduce all his characteristics. The brain scan can be analysed objectively to judge whether the replica is good enough to meet the agreed standards for personal identity continuation. If not, that person has died and someone else has been produced. Similarly, agreed standards can support the administrative and legal systems in accommodating cases of 'fission' and 'fusion'. There can be objective means for dealing with such problems as names, property distribution, rewards and punishments.

This thought experiment assumes a 'pure' physicalist account according to which the cerebral activity representing personal identity is capable of empirical investigation. If personal identity is an abstract 'entity' not totally manifested physically, as a possible interpretation of the Nozick-Cooper 'further fact' theory would have it (see chapter three), it will not be open to direct empirical investigation. There will have to be other means of dealing with disputes.

Personal concerns

The personal concerns of someone who has come to a conclusion about the metaphysical account of the self can be expected to reflect his beliefs. I have suggested that belief in an immortal soul may help someone to disregard worldly considerations. Normally, however, there is a similarity between individuals' personal concerns despite their diversity of beliefs. We nearly all want to be happy, healthy, good. To the extent that these concerns are derived from instincts, such as self-protection and the satisfaction of physical needs, they are an innate part of the individual's metric. To the extent that the concerns are learnt from

those around us they are social constructs. Even our basic instincts have evolved in such a way that we try to satisfy them in a social context.

Continuity

Perhaps more controversial is the claim that continuity as applied to personal identity cannot be fully realized except in social terms. Both rival accounts – biological continuity (in Olson's sense) and psychological continuity – seem to be applicable to the individual isolated from society. Imagine that the individual in the uninhabited world clones himself and removes the cerebrum from the cloned body. He then constructs a machine which removes his own cerebrum and transplants it to the cloned body. Afterwards the machine attaches the original body (containing the original brain-stem) to a life-support mechanism which maintains the biological system. In a way, the individual has achieved both biological and psychological continuity. But Mead would argue (see chapter two) that the creature had no self to begin with and has not achieved continuity of personal identity. I agree with that.

'I'

One putative advantage for the creature in the uninhabited world is that he does not need to communicate. He therefore does not need the word 'I' or any other words. The rest of us need to be able to refer to ourselves, and the word 'I', or its equivalent, is essential for self-reference. However, its use does not mean that there is some special referent associated with the word signifying more than is implied by the particular use made of it on a particular occasion. In self-consciousness subject and object merge, and sentences beginning 'I am....' do not describe a separate subject and object. This is worth saying because the very word seems sometimes to encourage people to speak or behave as if the referent of 'I' were a 'further fact'. In Jill Bolte Taylor's 'My Stroke of Insight' she describes how the disabling of one side of her brain and her eventual recovery seemed to enable her to choose which parts of her brain to use and, in that way, to

take control of her own personality. The stroke permitted her 'to choose who and how I want to be in this world'. (Bolte Taylor, 2008, p 122). She writes –and seems at times to have felt – as if 'she' were somewhere outside her brain operating it. This, of course, does not offend against common usage in such expressions as 'I pulled myself together' or 'I kept it in mind'.   Nevertheless, it is a deceptive use of 'I'.

My claim is that the linguistic significance of the word 'I' and of self-reference in general is sometimes conflated with metaphysical significance. Nozick, in his analysis of 'reflective self-reference', seems to encourage this confusion.   He suggests that 'the person is referred to in virtue of the property he acquires in the very act of referring' (1981, p 75). Yet the act of referring does not change the referent. Admittedly, concepts such as 'self', 'I', and 'personal identity' are strongly tied to the terms used to describe them, so that we learn how to use the terms as we learn about the concepts to which they refer. There must be a time when a child first says 'I', and a time – presumably separate and later – when he refers to himself 'reflexively' with the high level of self-awareness that Nozick associates with the acquisition of personal identity. The idea is comparable to James's suggestion that the onset of the self comes when we start to 'think of ourselves as thinkers' (see chapter 3). But, just as there must be a period between starting to think and starting to think about ourselves as thinkers, there must also be a period between starting to say 'I' and starting to use the word reflexively in Nozick's sense. Even if this first reflexive use of the word is distinguishable from all the previous uses, it is doubtful whether the pronouncement actually causes the change. It is not an *illocutionary* act such as the declaration of a marriage which brings about the actual event.

The concept
As with the word 'I', so with the concepts at least of 'self' and to a less extent, 'personal identity', there is an everyday use of which to take account. The

81

everyday concepts associate one person with one lifetime. (Although there are chinks in the armour of the everyday account. We talk of an individual not being the 'same person' as he was and of 'split personalities'.) Philosophers have over the centuries demonstrated that this everyday account will not do for them. The philosophical concept has been loosened by considerations such as fission, fusion and replication in general.

Parfit is right to suggest that, in this context, personal identity cannot be treated as depending on numeric identity. (See the discussion of 'U', the uniqueness factor, in chapter three.) If 'Am I about to die?' does not become an 'empty' question, it might take on a different meaning. 'Will my replication fail?' could express a reasonable worry. The cautious individual will postpone his assisted suicide until he obtains the result of his replication. Philosophers, free to suggest how the loose concept might be fastened down, are perhaps not able to fasten it, at least until practical scientific methods of replicating persons and investigating the results have caught up with theoretical speculation. Recalibration of the concept in order to accommodate replication might take into account some of the realities which do not seem well served by the everyday 'one lifetime' concept. Parfit, as I mentioned in chapter three, considers it 'natural' to apply the idea of 'successive selves'. Lewis (see chapter three) uses the phrase 'person stage'. We could apply the 'stage' notion to an ordinary lifetime and discuss what differences in personality could be treated as different 'person stages'.

At all times in dealing with the concept we need to take account of the social elements of personal identity and the social requirements the ideal 'fastened' concept must meet. Indeed, there is a sense in which society has the last word. Kafka's Gregor will have little time to enjoy his new person-stage as 'Beetle' if his family fails to recognize it as a new stage and crushes him. Cooper's *An Eldritch Tale* (2008a) does, I suggest, hold a key to how the concept can be developed. To understand the emergence of personal identity, the inter-dependence of 'individual authenticity' and 'social approbation' need to be

recognized. To 'fasten' a new concept a compromise is required, accommodating convergence of 'the individualistic and communitarian forms of relativism about the self' (p 136).

References

Barsalou, L.W., Simmons, W.K., Barbey, A.K., & Wilson, C.D. Grounding conceptual knowledge in modality specific systems. *Trends in Cognitive Sciences,* 7*,* 84-91. doi:10.1016/S1364-6613(02)00029-3

Cooper, W. (2002). *The unity of William James's thought*. Nashville: Vanderbilt University Press.

Cooper, W. (2008a). An eldritch tale: Kafka's *The Metamorphosis* and the self. *Philo, 11* (2), 133-144.

Cooper, W. (2008b). Nozick, Parfit, and Platonic glasses. *Sorites 20*, 98-116.

Descartes, R. (1993). *Meditations on first philosophy*. (D.A.Cress Trans.). Indianapolis: Hackett Publishing Company. (Original work published 1641).

Fichte, J G. (1994). *An Attempt at a New Presentation of the 'Wissenschafstlehre'*. (D.Breazeale, Trans. & ed.). Indianapolis: Hackett Publishing Company. (Original work published 1797/1798).

Gendler, T.S. (2009). Personal identity and metaphysics. In B.P.McLaughlin (Ed.), *The Oxford handbook of philosophy of mind* (pp. 580 – 591). Oxford: Clarendon Press.

Hume, D. (1911). *A treatise of human nature*. (A.D.Lindsay Ed.). London: J.M.Dent & Sons Ltd. (Original work published 1738).

James, W. (1947). *Essays in radical empiricism.* New York: Longmans Green. (Originally publishes 1876).

James, W. (1890). *The principles of psychology*. New York: Dover Publications.

James, W. (1891). *The moral philosopher and the moral life.* (Retrieved from http://www.philosophy.uncc.edu/mleldrid/American/mp&ml.htm).

James, W. (1982). An interview: Pragmatism – what it is. In H.S.Thayer (Ed.), *Pragmatism: the classic writings*. Indeanapolis: Hackett Publishing Co. (Original work published 1907)

Johnston, M. (1987). Human beings. *Journal of Philosophy, 84* (2), 59 – 83.

Johnston, M. (1989a). Dispositional theories of value. *Proceedings of the Aristotelian Society*, 63, 139 – 174.

Johnston, M. (1989b). Relativism and the self. In M.Krausz (Ed.), *Relativism: interpretation and confrontation* (pp. 441 – 472). Notre Dame: University of Notre Dame Press.

Kafka, F. (2009). *The metamorphosis*. (I.Johnston, Trans.). Retrieved from http://records.viu.ca/~johnstoi/stories/kafka-e.htm. (Original work published 1915).

Kant, I. (1977) *Prolegomena to any future metaphysics.* (P.Carus & J.W.Ellington Trans.). Indianapolis: Hackett Publishing Co. (Original work published 1783).

Lewis, D. (1976). Survival and identity. In A.O.Rorty (Ed.), *The identities of persons* (pp. 17 – 40). Berkeley: University of California Press.

Locke, J. (1975). *An essay concerning human understanding*. (P.H. Nidditch, Ed.). Oxford: Clarendon Press. (Original work published 1689).

Mead, G.H. (1934). *Mind, self, and society*. Chicago: The University of Chicago Press.

Nagel, T (1974). What is it like to be a bat? *Philosophical Review*, 83 (4), 435-450.

Nozick, R. (1981). *Philosophical explanations*. Cambridge, MA: Harvard University Press.

Olson, E.T. (1999). *The human animal – Personal identity without psychology*. New York: Oxford University Press.

Parfit, D. (1971). Personal identity. *Philosophical Review, 80* (1), 3 – 27.

Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.

Schelling, F.W.J. von. (1980). Of the I as principle of philosophy. In F.Marti (Trans.), *The unconditional in human knowledge*. Lewisburg: Bucknell University Press. (Original work published 1795).

Shoemaker, S & Swinburne, R. (1984), *Personal Identity*. Oxford: Blackwell.

Strawson, G. (2009). The self. In B.P.McLaughlin (Ed.), *The Oxford handbook of philosophy of mind* (pp. 541 – 564). Oxford: Clarendon Press.

Taylor, R. (2007). On the origin of good and evil. In L.P.Pojman & L.Vaughn, *The moral life* (pp. 141 – 153). New York: Oxford University Press. (Original work published 1984).

Williams, B. (1976). The self and the future. In J. Glover (Ed.), *The Philosophy of Mind.* Oxford: Oxford University Press. (Original work published 1970)